Endeavour Energy

Addressing the benchmarking factor for capex and opex

May 2014

Addressing the benchmarking factor for capex and opex

May 2014

Contents

About this report

Endeavour Energy has prepared an attachment as part of its regulatory proposal which demonstrates how it has met the objectives, criteria and factors in the Rules for capex and opex.[1] In the attachment, we set out how the criteria and factors should be considered as part of the AER's constituent decisions for opex and capex, and set out our evidence in relation to each factor. The purpose of this report is to provide further supporting information in relation to how we have addressed the benchmarking factor in the Rules.

Benchmarking is one of many factors the AER has to consider in making its decisions. The Rules require that the AER must:

> " …. consider the most recent annual benchmarking report that has been published under rule 6.27[2] and the benchmark capital expenditure that would be incurred by an efficient Distribution Network Service Provider over the relevant regulatory control period. "

Benchmarking is an undefined term in the Rules and can encompass many dimensions. The Productivity Commission has defined the term as any method for comparing a firm to other businesses, to itself over time (or between its various divisions) or to an ideal firm.[3]

We note that some of the measures of benchmarking our performance over time have been outlined in our responses to other factors in the Rules. In particular, we consider that an important benchmark is our performance against the AER's allowance in the previous period. Under the AER's incentive schemes, DNSPs are provided incentives to reduce expenditure levels below the targets set by the AER, and share these benefits with customers. Our response to these inter-related benchmarking factors are set out in the attachment that addressed the criteria and factors.

Key analysis and findings

The report examines the inherent limitations of benchmarking Australian DNSPs, and the role that benchmarking should play as a partial indicator of efficiency. Our analysis identifies that benchmarking has inherent limitations such as inability to conduct 'like for like' analysis across peer firms, data inconsistency and inaccuracy, and failure to meet statistical principles. We think that valid benchmarking may have a role in guiding the regulator to areas requiring further granular analysis. It should not be used to reject a DNSP's proposal, or as a basis to substitute the forecast given the inherent limitations as a tool.

The report assesses the relative weight that should be applied to each of the benchmarking tools identified by the AER in its Forecast Expenditure Assessment Guidelines including economic analysis, aggregated category analysis, and cost category data including the augex and repex models. When deciding if a benchmark is appropriate, we have been guided by the Productivity Commission's review in 2013 which set out 6 criteria for when a benchmarking tool could be used in the process. This includes validity, accuracy and reliability, robustness, simplicity, not subject to manipulation and fitness for purpose. To complement this analysis, we have also sought to understand the available data that can be used for benchmarking and reported on these outcomes. This was based on a detailed Huegin Consulting study of 7 DNSPs in Australia, and data of other DNSPs where available.

Based on this assessment approach, we have placed limited weight on benchmarking analysis as a valid test of the efficiency of our forecast and consider that the AER should do likewise in its assessment. In all cases, the AER's techniques do not meet all the criteria specified by the Productivity Commission. In some cases, such as economic analysis we consider the method may actually provide misleading results and should not be used by a business or the AER to test efficiency. In other cases, the model may provide some insight into the efficiency of a DNSP's forecasts, for instance when the data quality is sound. In these cases, we have considered the underlying data and provided commentary on any observed differences in light of our circumstances and drivers of expenditure.

Our analysis of benchmarking tools suggests that trends in a DNSP's results over time are of more value, than relative efficiencies between DNSPs at a point in time. In this respect the data provided does demonstrate that Endeavour Energy's growth rates in expenditure are among the lowest out of the peer group studies. Once again, we draw caution on such results as they cannot capture the reasons for observed differences between DNSPs.

[1] Attachment 0.03 "Endeavour Energy: Addressing the Objectives, Criteria and Factors for capex and opex in the NER", May 2014.
[2] The AER intends to release its first benchmarking report in September 2014, and therefore we are not provided with an opportunity to demonstrate or make representations on this report at the time of submitting our regulatory proposal.
[3] Productivity Commission, *Electricity Network Regulatory Frameworks*, 26 June 2013, p147.

Structure and contents of document

We have structured our response as follows:

- Section 1 identifies the inherent limitations of benchmarking data and the role of benchmarking. We provide an outline of how we have assessed each of the AER's benchmarking tools outlined in its Forecast Expenditure Assessment Guidelines.
- Sections 2 sets out our findings on economic analysis benchmarking.
- Section 3 sets out our findings on aggregated cost benchmarking.
- Section 4 sets out our findings on cost category comparisons. Attachment 5.07 and 5.08 provide further information on a joint analysis undertaken by NSW DNSPs on the effectiveness of the augex and repex models respectively.

In this document we have referred to a series of reference material.

Attachment number	Name of document
Attachment 0.11	The Huegin report for Endeavour
Attachment 5.07	Joint DNSP analysis on effectiveness of the augex model
Attachment 5.08	Joint DNSP analysis on effectiveness of the repex model

1 Limitations and role of benchmarking

When addressing the capex and opex criteria in the Rules, the AER must consider all the factors in the Rules, including benchmarking. This means that the AER must come to a view on the extent to which benchmarking is relevant to its constituent decisions in respect of capex and opex, and the weight that should be applied to the analysis it examines. This requires the AER to consider the extent of information available and the probative value of the analysis.

The purpose of this chapter is to provide context on the role of benchmarking as a tool in providing a 'partial indicator' on the efficiency of a DNSP's forecast.

- We demonstrate that benchmarking data has inherent limitations, which means that extreme caution must be applied in using the analysis to draw conclusions on the relative efficiency of a DNSP.

- With this in mind, we demonstrate that benchmarking can play a role in helping the regulator or a business uncover potential inefficiencies in its forecasts, but that the data cannot be used directly to reject or substitute the proposed expenditure of a DNSP. Rather the analysis can be used to target the AER's analysis.

- We conclude the chapter by setting out our method for reviewing the AER's benchmarking techniques in the Forecast Expenditure Assessment Guidelines.

1.1 Limitations of benchmarking

We consider that benchmarking data and techniques should be used with extreme caution due to the errors that arise when using such tools to measure efficiency. The errors arise from the inability of models to normalize for the inherent differences between DNSPs, data inconsistency and inaccuracy, and low statistical reliability.

The primary issue with benchmarking is that electricity distributors are not homogonous in the Australian market, and therefore 'like for like' comparisons cannot be used effectively to draw inferences on efficiency. For example:

- Each DNSP in Australia operates under unique conditions such as customer density, geographic area, topographic conditions, and the inherent design of network (for instance number of sub-transmission assets). This makes 'like for like' comparisons highly problematic, as there is no statistically reliable method for normalising data. For example, a rural DNSP may perform better if expenditure data is normalised by line length, but will likely perform worse if the data is normalised on a per customer basis.

- DNSPs are at relatively different stages of the investment life cycle. This impacts heavily on relative replacement and maintenance costs, with older networks incurring a higher inherent cost of performing its functions relative to a younger network.

- Jurisdictional differences also play a major part in explaining differentials in cost structures. For example, jurisdictions differ in respect of licence conditions, markets for contestable services, and classification of services.

In addition to these issues, benchmarking is often plagued by data issues. Up to this point in time, there has been limited granular data on DNSP's expenditures and operating environment. Even when data is available, we note that it rarely is in a consistent form to draw meaningful judgments. For example, cost accounting policies are so varied between DNSPs that they do not even allow for effective comparison of opex and capex between DNSPs. This means that limited weight can be applied to this data when using it for the purposes of identifying efficiency at the high level.

Finally, we note that our experience with models used by the AER is that they largely fail to meet key principles of statistical validity such as:

- Sample size – In the absence of a large sample size, the results can be skewed by a few firms with similar outcomes. Further, low sample size is counter-intuitively related to the apparent strength of relationship that is observed in tests such as R squared, leading to misleading conclusions. In Australia, there are an insufficient number of firms to develop a large enough sample size.

- Comparability of data series – Even within the small sample of Australian DNSPs, there is a great degree of variability in the underlying conditions of distributors in Australia. This breaches a key test of statistical validity which requires the assumptions underpinning data inputs to be consistent.

- Correlation between dependent variables – In multi-variable models such as Total Factor Productivity, many of the factors are highly inter-related with each other. For example, SAIDI and SAIFI will be highly related to each other.

- Consistency in variations across data range –Tests such as R squared can be misleading if the data series does not show a consistent error range across the data series. This underscores once again the importance of a large sample size.

- Sensitivity analysis does not lead to wide variation in outcomes – An important aspect of statistical application is to test whether the outcome is relatively consistent when other likely variables are used in the analysis. For example, energy consumption and peak demand are two types of outputs delivered by a DNSP.

For these reasons, we have applied a high degree of caution in interpreting benchmarking analysis when testing the validity of our forecasts. In particular, our experience is that high level benchmarks serve to validate the inherent differences between DNSPs, or reflect data inconsistency and incomparability.

The limitations of benchmarking have been emphasised in a review by the Productivity Commission. In 2013, the Productivity Commission was commissioned by the Australian Government to review the extent to which benchmarking could be used by regulators in the electricity industry. The Productivity Commission directly referred to academic articles on the need to select explanatory variables that describe the functions undertaken by a DNSP and the environment in which it performs. In particular it referred to an article by Turvey in 2008 which stated:[4]

> "Comparisons between networks of the costs of these activities can only illuminate differences in the efficiency with which operations and maintenance are carried out if the magnitudes of the tasks of operation and maintenance can be compared. This is a platitude, yet failure to articulate it has led some authors to scrabble around among available data to select a set of "explanatory" variables without displaying any understanding of what an enterprise does and how it does it. Confusion about these matters is rife, as witnessed, for example, by the fact that while some econometric efficiency estimates for electricity distribution treat MWh distributed, km of overhead lines or number of customers as an input, others treat one or more of these variables as an output."

The ACCC's comprehensive review of benchmarking capex and opex in energy networks also came to similar conclusions as the Productivity Commission. The ACCC observed that all benchmarking results in a degree of approximation due to its abstract nature in aggregating many factors into a few variables. It stated[5]:

> "Effective benchmarking requires the modelling of relevant factors affecting the expenditure of the energy networks. These businesses provide a range of services using different types of inputs and may operate in different environmental conditions. Inevitably, benchmarking requires some aggregation of those services, inputs, or environmental conditions into a few variables, resulting in some degree of approximation in the estimation."

1.2 Role of benchmarking as a partial indicator

Despite the limitations identified above, we consider that well designed tools can play a role for a business or regulator to test the efficiency of a forecast. In the sections below, we identify two principles in applying benchmarking information:

- Testing the relative effectiveness of the tool to provide insights into efficiency.
- Using the information in a way that enables the decision maker to identify whether there is a potential inefficiency in the forecasts

Effectiveness of benchmarking

Benchmarking tools need to be designed in a way that provide insights into potential areas of inefficiency, Poorly designed tools or misapplication can mislead the decision maker, resulting in outcomes that do not meet the long term interests of customers. This was a key finding of the Productivity Commission[6]:

> "A key question is how to separate the wheat from the chaff among the various competing approaches, recognising that this will typically involve balancing various criteria …. There is a large literature on estimating the comparative costs of businesses, with much of that literature concentrating on using the 'right' techniques. However, it is equally important to be clear about how to interpret benchmarking results for policy purposes because the misuse of good technical analysis can result in adverse outcomes for consumers and businesses. In particular, comparing the costs between businesses in different jurisdictions without accounting for factors outside the control of the business could provide misleading indicators of managerial efficiency. If used in incentive regulation, this could lead to underinvestment or unwarranted transfers from consumers to the businesses."

The Productivity Commission set out 6 formal criteria for identifying the effectiveness of benchmarking. These include[7]:

---

[4] Productivity Commission, *Electricity Network Regulatory Frameworks*, 26 June 2013, p160.
[5] Australian Competition and Consumer Commission, Benchmarking Opex and Capex in Energy Networks Working Paper no.6, May 2012, p12
[6] Productivity Commission, *Electricity Network Regulatory Frameworks*, 26 June 2013, p163.

- Validity - A valid benchmark should relate to efficiency (or conversely inefficiency) in one or more meaningful dimensions. A valid benchmark should reflect the way that the businesses are run. In particular, comparing the costs between businesses in different jurisdictions without accounting for factors outside the control of the business could provide misleading indicators of managerial efficiency.

- Accuracy and reliability - Accuracy is the degree to which a benchmark provides an unbiased estimate of efficiency, while the reliability (used here in the normal sense of reproducibility) is about the variance of the measure.

- Robustness - This is a subset of accuracy and reliability, but worth emphasising in its own right. A particularly useful robust measure is one that provides information about the efficiency of an enterprise regardless of its operating environment.

- Manipulation and gaming of data - As in all systems where rewards and punishments depend on incomplete measures of performance, the measured party has incentives to ‘look’ like a highly performing entity. Accordingly, the regulator should consider the capacity of any particular benchmarking measure to create unforeseen business behaviours.

- Parsimony - A good model should be no more complex than required. This is important in assisting interpretability, avoiding data mining, achieving robust results, reducing data collection costs and allowing greater comparability of results across countries.

- Fit for purpose - Benchmarking has multiple purposes. Some require great accuracy, reliability and robustness. This is particularly important where benchmarking is used to determine a business’s revenue allowance. Such benchmark estimates should be highly reliable across time, business types and jurisdictions. The concerns are less where benchmarking is indicative — used to identify areas for possible future investigation, or to reach some prima facie judgment.

The Productivity Commission’s criteria set a very high threshold for an effective benchmark, and it is unlikely that any tool would meet all the criteria given the inherent limitations of benchmarking. Given this is the case, there needs to be an element of subjective judgment in deciding whether to disregard a tool in entirety, or whether it may still have some probative value as a partial indicator of efficiency. We consider that some tools may actually lead the decision maker into error, and cannot be verified by reference to a review of programs or projects.

Using benchmarking data to assess the efficiency of a forecast

In some cases, a tool may satisfy some of the criteria of the Productivity Commission’s criteria, and be suitable to be used as a partial indicator of efficiency. This does not mean that the tool can be used in a deterministic way to reject proposed expenditure, or to develop alternative or substitute forecasts.

Rather, the analysis could be used as an informative tool to identify potential areas of inefficiency in a forecast, and to target reviews on these areas. This would then require granular analysis of the proposed programs or cost category to check whether the observed result relates to inefficiency, or stems from a reasonable driver of expenditure. This in turn requires a degree of expertise by the decision maker using the tool both from a statistical and engineering perspective. The ACCC came to this view when undertaking its review of capex and opex benchmarking:

> “Reflecting current practice and existing expertise, benchmarking should initially be used as an informative tool rather than a determinative one. For example, it can be used as a starting point for a conversation with regulated utilities about the level of operating and/or capital expenditures being incurred and proposed. A more sophisticated application could emerge over time.
>
> Effective cost benchmarking requires a clear understanding of the structure of the costs of the regulated utilities. This, in turn, requires an understanding of the key outputs provided by the benchmarked utilities, the inputs used (and/or the prices of those inputs), and the key environmental factors. It is also useful to understand the nature of any economies of scale or scope in the industry. Engineering studies can help provide a picture of the likely cost drivers, including how the cost drivers interact. This involves complementing in-house resources through access to expert consultants with specialised engineering knowledge and experience in the application of cost-benchmarking methods.”[8]

---

[7] This is based on discussion in the Productivity Commission’s report, *Electricity Network Regulatory Frameworks*, 26 June 2013, pp168-186.

[8] Australian Competition and Consumer Commission, Benchmarking Opex and Capex in Energy Networks Working Paper no.6, May 2012, p14

1.3 Analysis of available benchmarking data and models

We have applied the principles described in Section 1.2 above to the methods identified by the AER in the Forecast Assessment Expenditure Guidelines. We have sought to understand the probative value of these tools using the formalised criteria developed by the Productivity Commission. We have also commissioned reports by Huegin Consulting which helped us to understand the data available at this point in time in respect of applying these tools, and reported on the results. In the following chapters we describe our findings in relation to:

- Economic analysis;
- Aggregated category benchmarking; and
- Cost category benchmarking including the repex and augex models.

We recognise that over time we will have access to better industry data, together with a better understanding of the factors that drive cost differences between DNSPs. The AER is currently undertaking its first information gathering exercise for the purpose of producing a benchmarking report. We consider that increased data does not necessarily improve the probative value of benchmarking tools, and that at all times the decision maker needs to consider whether the tool is effective, and can be used to support more granular analysis of programs and projects.

2. Economic benchmarking analysis

The AER has defined economic benchmarking as applying economic theory to measure the efficiency of a NSP's use of inputs to produce outputs, having regard to environmental factors. There are a number of methods and tools to undertake economic benchmarking such as Total Factor Productivity (TFP)[9] and Data Envelopment Analysis (DEA).[10]

2.1 Effectiveness of technique in guiding decision making

Previous concerns on effectiveness of economic analysis

In late 2013, the AER issued a Regulatory Information Notice to Endeavour Energy and other DNSPs to collect information relevant to economic benchmarking analysis. In our response to the AER's draft RIN on economic benchmarking, we noted that the application of these tools to guide regulatory decision making would result in error, leading to outcomes that are detrimental to the long term interests of customers. Our view was based on the following reasons:

- We are not convinced that economic benchmarking tools such as Total Factor Productivity (TFP) can be used to infer relative efficiency of DNSPs over time. We consider that the models cannot adequately normalise for differences between DNSPs, and do not provide meaningful assessment of the apparent differences in productivity levels. For example, TFP will show that a firm that replaces ageing assets has declining levels of capital productivity, as the model would show higher prices for capital while maintaining existing service levels. In our view this would be driven by the age of the asset base which is likely to vary between DNSPs.

- We consider that economic benchmarking models such as TFP do not provide the AER with guidance on how to target its review of expenditure forecasts, as the information provided is at too high a level to identify potential areas of efficiency. The models and data collected will not provide any guidance on the underlying drivers of apparent productivity rates, and therefore does not provide useful analysis that identifies which areas to review in a DNSP's capex and opex forecasts.

- Finally, the DNSPs in Australia have repeatedly noted that our finance and asset systems have not recorded much of the data in the form required by the AER, and that the information used to populate models would be highly unreliable. This has been independently confirmed by the Australian Energy Market Commission, who conducted a 2 year review of TFP and concluded that available historical data was of poor quality and reliability.

We provided the information required by the RIN to the AER on 30th April 2014. Our basis of preparation has noted areas where we have had to provide best estimates when data was not available in our systems. We have also noted instances where the definitions could be open to interpretation. We understand that this has also been the experience of other DNSPs, which draws into question the quality of information that would be used as part of economic benchmarking. In this respect, our experience under industry reform has also found significant differences between the 3 NSW DNSPs, which would not have been identified if further investigation had not occurred.

This is very concerning when models such as TFP and DEA rely on the totality of variables to form rankings and assessments of efficiency over time. An error in one variable can lead to significant deviations in observed performance. When these errors are multiplied across many variables, the outcomes of the analysis could not be used to infer efficiency.

In addition to data quality issues, we note that there are a number of model specifications being considered by the AER to undertake economic benchmarking.[11] The variety of model specifications shows the difficulty of deriving an input and output relationship that can adequately account for the nature of the industry, and the inherent differences between DNSPs. For example, the outputs generally used are energy consumption, peak demand supplied, capacity of the

9 The AER noted that this when benchmarking businesses that have more than one output and/or more than one input the challenge is including these different values into a common comparable index. MTFP uses revenue and cost shares as weights to overcome this problem and create a value for a firm's output and a value for a firm's input; the productivity of the benchmarked firm is then the difference between these output and input figures. In the context of Australian NSPs this means that a businesses efficiency will be affected by its outputs compared to the industry average and the share of expenditure that these outputs account for.

10 The AER note that this is a more limited technique than MTFP, because it cannot incorporate as many input and output variables and because it requires more data. Therefore, it proposed using it to cross-check the results of the MTFP analysis. It may be possible to decompose the efficiency scores of DEA to identify different types of inefficiency.

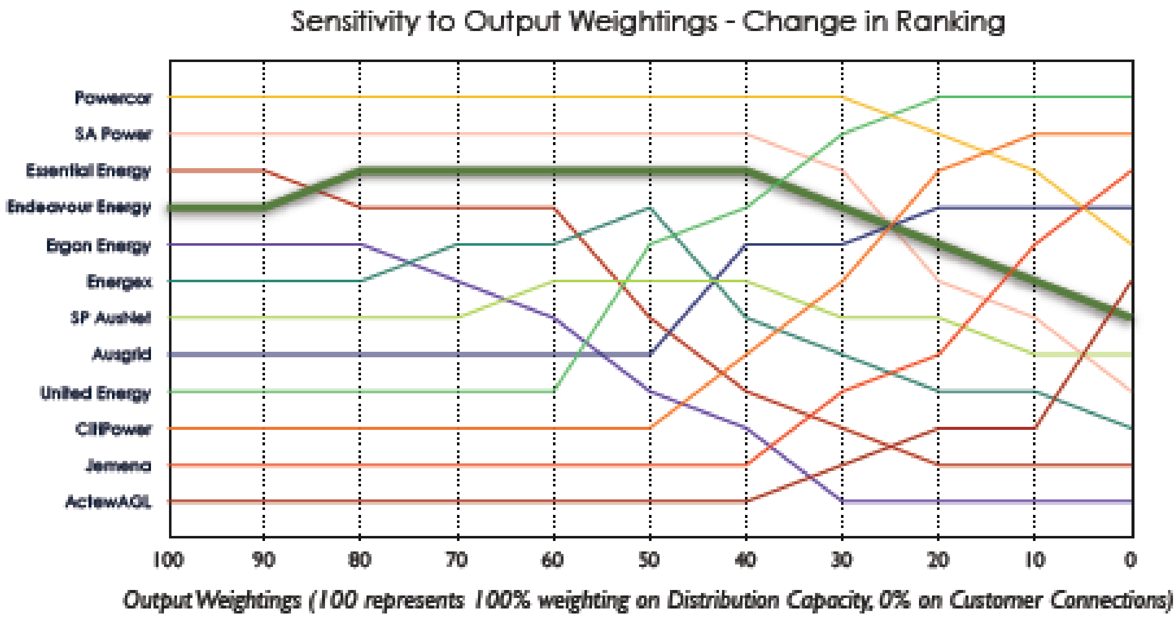
11 We note that the AER has yet to provide DNSPs with its model specifications, meaning that we are unable to undertake meaningful analysis on whether the model specification is appropriate.

network and reliability. These output variables are highly integrated, and often will have no relationship to input costs over time. A very simple example is when a DNSP has to undertake significant replacement of its network. While its costs increase, it is producing the same level of output[12] and therefore shows a decline in productivity.

The lack of a precise model specification means that there is a high degree of subjectively in the model applied by the analyst. This point has been borne out in the evidence provided by Huegin Consulting. Huegin noted that[13]:

> "Model errors and bias are always present in economic analysis, which is not an issue in itself, but the diversity of conditions in Australia and in the inherent network designs means that a particular model specification will provide advantage for some businesses and disadvantages for others."

Huegin demonstrate this point by undertaking its own economic benchmarking models. The following diagram shows that variable selection and weightings can skew the outcomes of models such as TFP. In the diagram it shows that Endeavour Energy ranks 4th among the DNSPs in Australia if distribution capacity is used as an output. In contrast, if customer connections are selected as the variable, then Endeavour Energy's ranking falls to 7th. Such sensitivity analysis demonstrates that economic benchmarking models reflect the underlying characteristics of the DNSP rather than suggest relative efficiency.



Based on the analysis above, we consider that the use of economic benchmarking should not be used by the AER as it fails to meet any of the criteria identified by the Productivity Commission for effective benchmarking.

- Validity – We consider that economic analysis is not a good determinant of relative efficiency of a DNSP at a point in time, and is more likely to reveal the underlying network and accounting differences of each of the businesses. This can be seen in the diagram above, which shows that a change in the weighting of output variables (only one variable among many) results in vastly different outcomes. Similarly, performance over time can be misleading as a change in the value of one variable may lead to significant changes in perceived efficiencies, even if this variable can be explained with reference to underlying circumstances. For example, the model could not explain when a DNSP is investing to meet a step change in the security of the network as a result of new licence conditions. In these circumstances the RAB would increase greatly, but there may not be a direct relationship with output factors such as reliability, although there may be some improvement in capacity of the network.

- Accuracy and reliability – We consider that the underlying data provided by Endeavour Energy and other DNSPs is not reliable across all variables. This is concerning given that economic analysis is based on multi-variable models that are only statistically credible if each variable is correct. In developing the data requirements to undertake economic benchmarking, the AER required DNSPs to submit information that could not be provided from systems and financial statements. The AER required that DNSPs provide a 'best estimate' in these cases, and document the methodology. The best estimates were subject to a less onerous form of external assurance. Our own experience is that much of the information provided to the AER was derived from

[12] For instance, replacement does not increase capacity of the network. Replacement may improve reliability, but in most cases replacement is to maintain the existing level of reliability on the network (ie: stem a decline in reliability).
[13] Huegin Consulting, "Distribution benchmarking study: Endeavour Energy", 2014, p10.

estimates that cannot be verified, and therefore should not be used in economic analysis, which in itself is required to support changes to expenditure programs. We understand that other DNSPs have also submitted information derived from estimates, and therefore the data set upon which the models are based cannot be relied on by the AER.

- Robustness – We consider there are no robust model specifications that can be used for economic analysis. Huegin's analysis notes that all model specifications are subject to bias, and that the outcomes from economic analysis are highly sensitive to the model specification applied. This is particularly the case where there are a small number of DNSPs and they are highly heterogeneous.

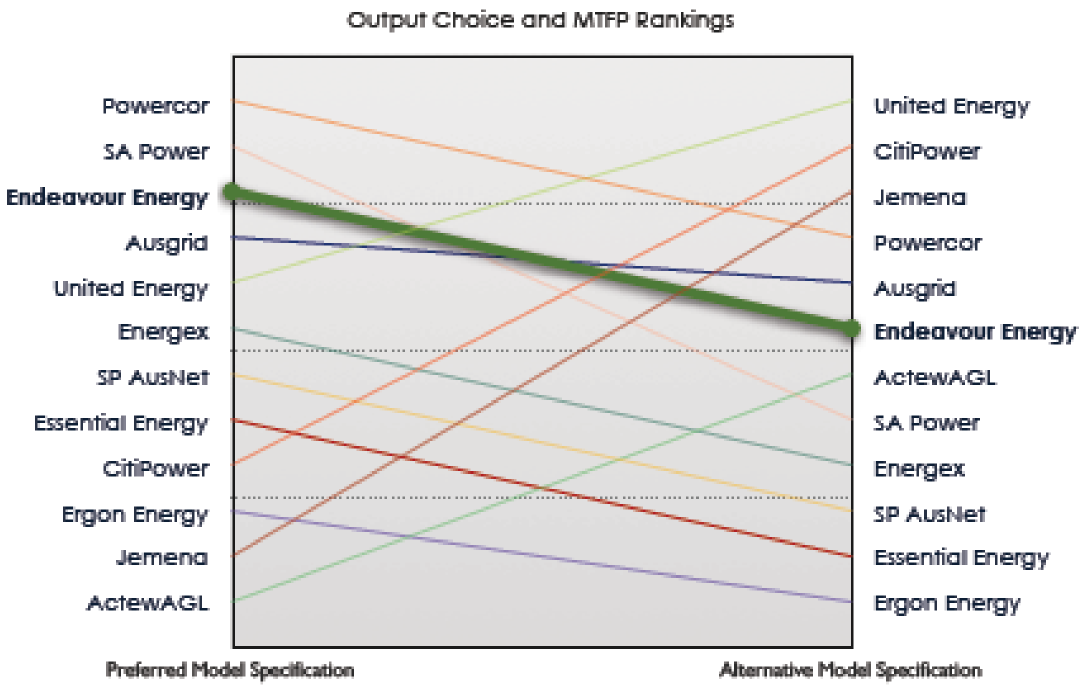
- Manipulation of data – The AER's requirement that DNSPs provide best estimates where actual data cannot be provided, together with ambiguity of definitions, may leave it open to a DNSP to report in a manner that puts its circumstances in the best light. This once again underscores the unreliability and accuracy of the information.

- Parsimony – In some respects, economic analysis is a simple tool to use, however the complexity of model specifications and the need for expert statistical application mean that the tool is not simply applied in practice. For example, the analyst needs a firm understanding of data quality and statistical principles when selecting the appropriate model design and drawing inferences from that method. At this stage, the AER has not identified the model specification it will use for economic benchmarking, which raises further questions over whether the tool to be applied will be simple.

- Fit for purpose – This is perhaps the deepest difficulty with economic analysis. Due to its lack of granularity, economic analysis is almost impossible to understand the drivers or potential inefficiency underscoring the outcomes. This means that it is an excessively poor tool for a business or a regulator to target areas of the program for further review. For instance, even if the analysis can be used to show that change in capex levels (ie: change in value of regulatory asset base) is the driver of perceived inefficiency, it does not provide the business or regulator with any information on which programs or projects should be targeted.

2.2 Analysis of available data

Based on these considerations, we consider that economic benchmarking should have a zero weight in the AER's decision making. If the AER still consider that it provides some form of guidance in its decision making, we note that Huegin's analysis suggests that Endeavour Energy performs 3rd best for the preferred model specification, and 6th for alternative model specifications.



We consider that these results do not provide a cohesive argument for suggesting that Endeavour Energy is relatively efficient relative to other DNSPs. We consider that this would be misleading and that the outcomes relate to the model specification selected by Huegin and the underlying network characteristics of Endeavour Energy in relation to the model variables and weightings.

3. Aggregated category benchmarking

Aggregated category benchmarking captures information such as how much a NSP spends per kilometre of line length or the amount of energy it delivers. This can be undertaken for capex or opex.

3.1 Effectiveness of technique in guiding decision making

Aggregated benchmarks have certain advantages over economic benchmarking techniques. A key advantage is that data on actual expenditure, and statistics such as line length, customer numbers, transformer capacity and square kilometers have a high degree of accuracy. It therefore meets the Productivity Commission's review of accuracy and reliability of data, and also limits the ability of DNSPs to manipulate the data in a favorable way. Further the simplicity of the tool in terms of regressing a single variable meets the criteria of parsimony.

The weakness of aggregated techniques is that they fail to meet the other important criteria of the Productivity Commission.

- Validity – There are many network and accounting drivers that underpin differences in the comparative data of DNSPs. They therefore form an important element of explaining relative cost differences, and it is almost impossible to identify whether any particular area of capex or opex is inefficient unless such costs are normalised or removed. This reflects that the tool is not sufficiently granular to identify where variations in costs are occurring and whether these relate to inherent inefficiencies.

- Robustness – Similar to economic analysis, aggregated category benchmarking is subject to model specification issues, and therefore it is difficult to form conclusive opinions on relative efficiency. For instance we note that normalising the data for line length, customer numbers, peak demand, or energy consumption can skew the outcomes. This shows that while DNSPs may not be in a position to manipulate the data, that the analyst is able to choose metrics that may provide support for a pre-conceived view.

- Fit for purpose – The difficulty with the approach is that it is undertaken at a very high level, and does not contain any additional information on where potential inefficiencies may lie. At worst, it can mislead the business or regulator into considering there are inefficiencies which may bias bottom up reviews of expenditure programs.

We consider that using aggregated benchmarks to infer the relative efficiency of DNSPs should be used with extreme caution, and with regard to all factors that may explain performance. Our view is that, if the information is used, it should be accompanied by detailed granular benchmarking of cost categories as discussed in Chapter 5.

While there are difficulties with inferring the relative efficiencies of DNSPs, we consider that aggregated benchmarks do provide insight on how a DNSP performs over time, such as growth rates in expenditure. The advantage of comparing growth rates is that it uses a consistent data series. This was noted by Huegin Consulting when it stated:[14]

> "Understanding where a business stands in the rankings of productivity of industry participants is interesting, but perhaps not useful. An understanding of the difference between modeled future costs and an individual business' forecasts is useful."

We consider that growth rates over time are useful for the AER's trend analysis, and to identify if there is a particular driver at play for explaining relative growth rates of DNSPs. For instance, at a macro level the AER may identify that a higher relative growth rate is related to a new licence condition, or change in replacement approach which requires further investigation. For a business, it also provides a macro marker of whether efficiency initiatives, such as those we have undertaken as part of industry reform in NSW, are yielding positive results in comparison to expenditure trends of other DNSPs. Caution needs to be attached to any firm conclusions drawn from this type of analysis, as there are questions of validity, robustness and fitness for purpose must be examined.

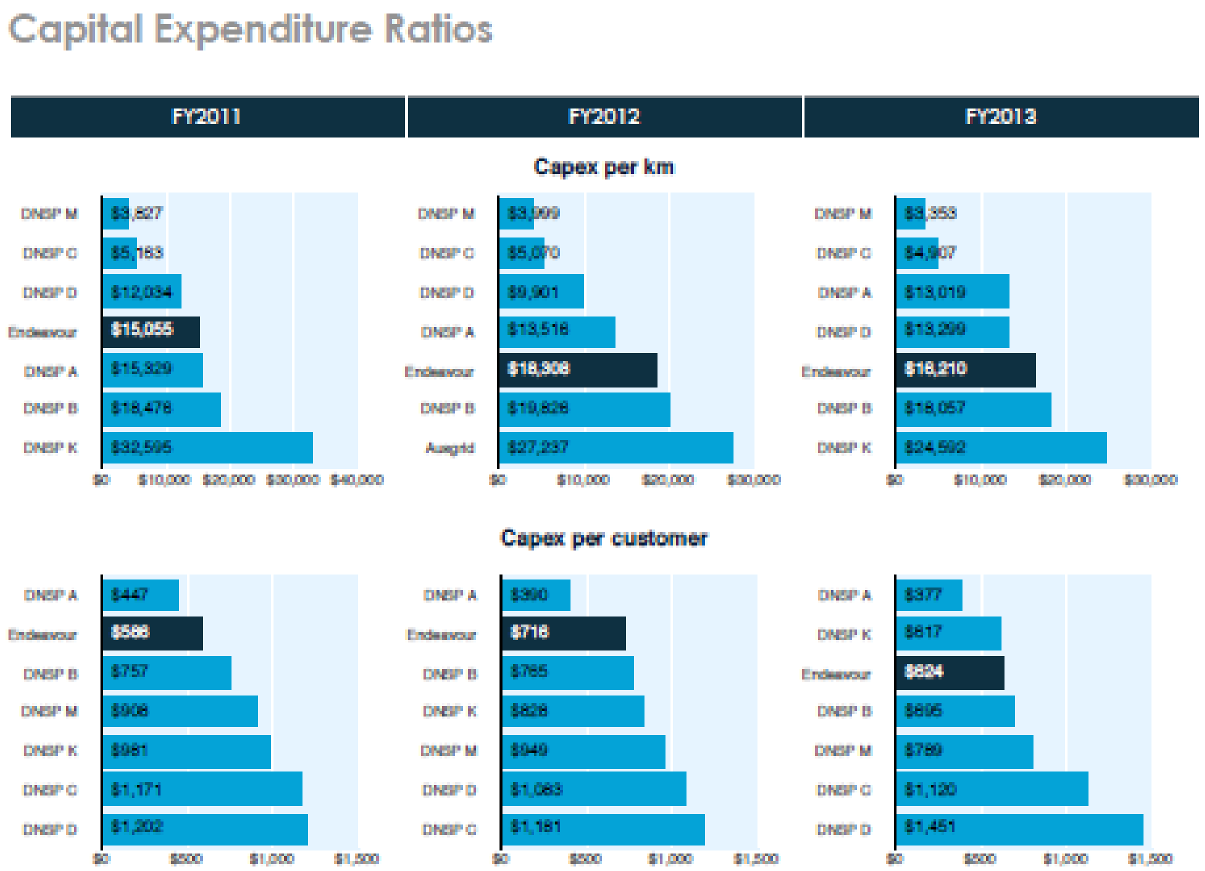
With this frame in mind, we sought to understand where Endeavour Energy fits with respect to different forms of aggregated benchmarks.

[14] Huegin Consulting, "Distribution benchmarking study: Endeavour Energy", 2014, p12.

3.2 Analysis of available data

Capex

The diagrams below show Endeavour Energy's capex per kilometer and by customer for each year from 2010-11 to 2012-13. The analysis underscores the weaknesses of using aggregated benchmarks in assessing our relative performance. For instance, on a kilometer of line basis we perform in the middle of the range, while on a capex per customer basis we perform close to best in the group. Once again this highlights that our inherent network characteristics as a predominantly urban DNSP, which nevertheless has a wide geographic area. The diagram shows that Endeavour Energy's capex on both these metrics rose between 2010-11 before falling again between 2011-12 and 2012-13. This reflects that our capex increased in 2011-12 but then has flattened since that time consistent with NSW industry reform.



More importantly, Endeavour's capex starts to fall markedly when the forecasts for 2014-19 are compared to actual capex in 2009-14. In the 2009-14 period, Endeavour Energy's compound annual growth rate was 5.1%. Endeavour Energy's growth rate is expected to decline over the 2014-19 period, with a compound annual growth rate of -7.3 per cent.

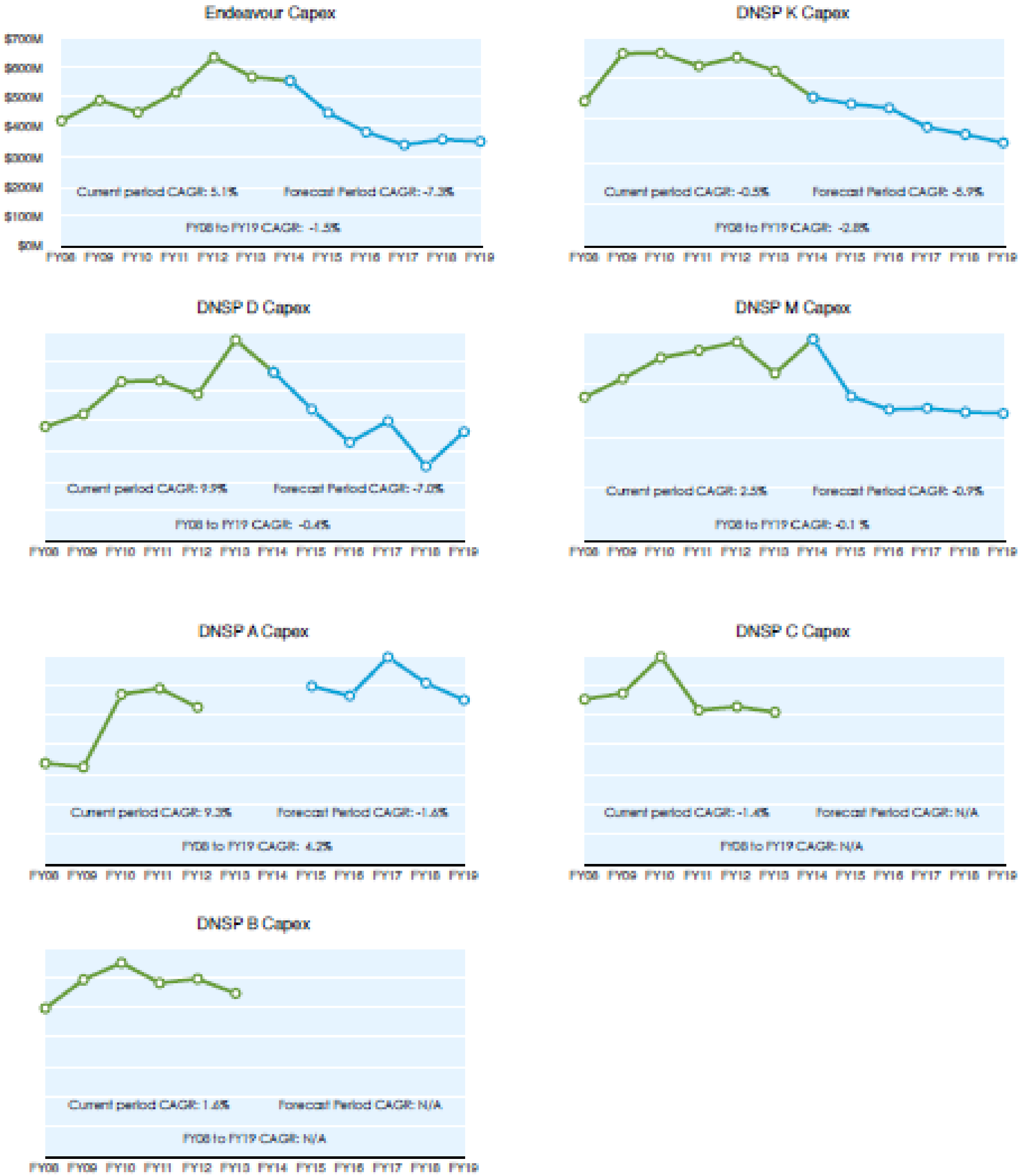
In our view, this analysis provides a high level view on the effectiveness of industry reforms in driving efficiency within our businesses. We believe that the evidence provided by Huegin provides a rough 'rule of thumb' to support the position that industry reform has delivered significant efficiencies in our capex forecasts. However, we consider the data by itself does not provide compelling data for the AER to draw sound conclusions on the efficiency of the forecast. This is for 3 reasons:

- Granular data would need to be reviewed to compare the underlying drivers of Endeavour Energy's capex trends, including granular assessment of the investment programs, drivers and processes. Despite the apparent reductions in capex, we consider that it is not possible to ascertain whether the profile is consistent with changes in our regulatory obligations, demand environment and condition of networks on our assets.

- Similarly trend data comparisons with other DNSPs cannot adequately account for circumstances driving investment in other DNSPs. For instance, other DNSPs may be undertaking similar efficiencies at the same time, or may have found a need to increase their replacement programs.

- We note that capex-opex substitution possibilities can impact the comparison over time. For instance, the reductions in capex may have an impact on opex forecasts.

DNSP Capex Trends - Actual and Forecast

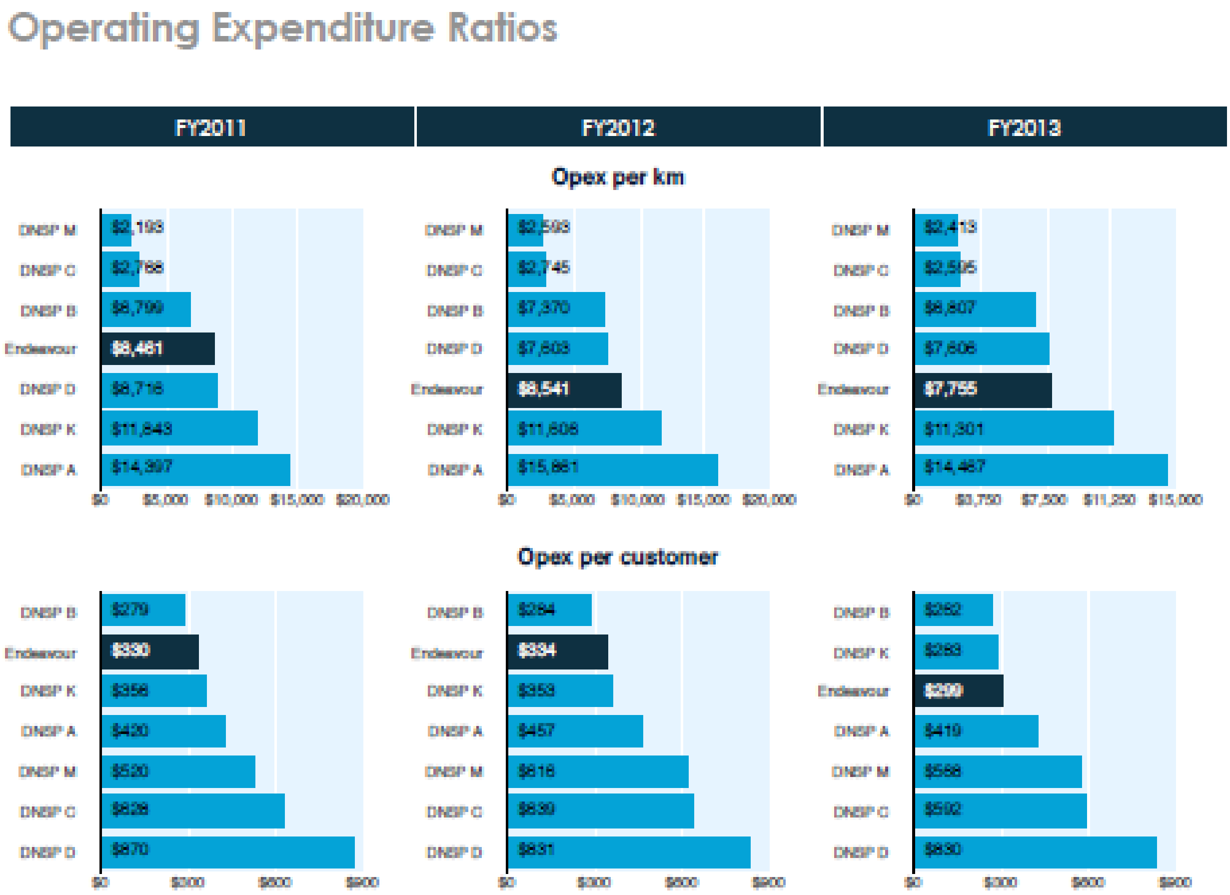


Notes:
1. All figures in FY2013 dollars.
2. Data sources are the RINs and directly supplied data from DNSPs.

Opex

The diagrams below show Endeavour Energy's opex per kilometer and by customer for each year from 2010-11 to 2012-13. Similar to capex, the analysis underscores the weaknesses of using aggregated benchmarks in assessing our relative performance. For instance, on a kilometer of line basis we perform in the middle of the range, but on an opex per customer basis we are among the best performers. Once again this highlights Endeavour's inherent network characteristics as predominantly an urban DNSP, which nevertheless operates across a relatively large geographic area.

The diagram shows that Endeavour Energy opex on a per kilometer and per customer basis increased between 2010-11 and 2011-12, but reduces in magnitude in 2012-13   Our relative ranking for these measures stayed relatively consistent.



Huegin has provided information which compares Endeavour Energy's opex levels over the current period (2009-14 period) through to the forecast period (2014-19) relative to other DNSPs. In the 2009-14 period, Endeavour Energy's compound annual growth rate was 1.2%, but will fall to -0.4% for the 2014-19 period.

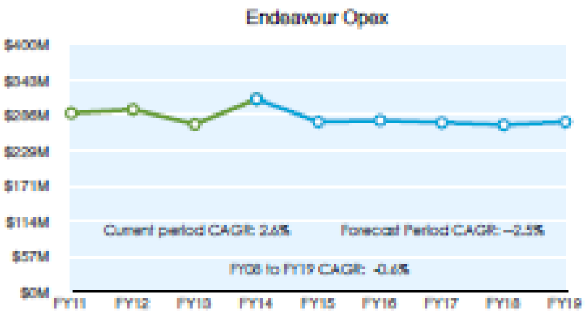
In our view, this was important in providing a high level view on the effectiveness of industry reforms on driving efficiency within our businesses. We believe that the evidence provided by Huegin provides a rough 'rule of thumb' to support the position that industry reform has delivered significant efficiencies in our opex forecasts. We note that trends in opex over time are likely to provide more information than trend data for capex which is lumpy in nature. Nevertheless comparisons need to consider:

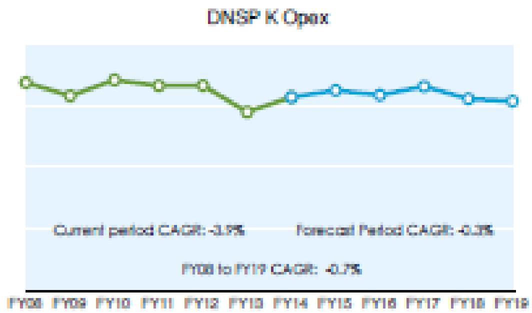
- The underlying drivers of Endeavour Energy's opex trends, including assessment of trends in cost categories. For example the trend may be affected by deterioration or improvement in underlying asset condition, which in turn influences maintenance costs. Similarly, the net benefit of efficiency programs (ie: the savings minus the costs) may be influencing costs in particular categories of expenditure.

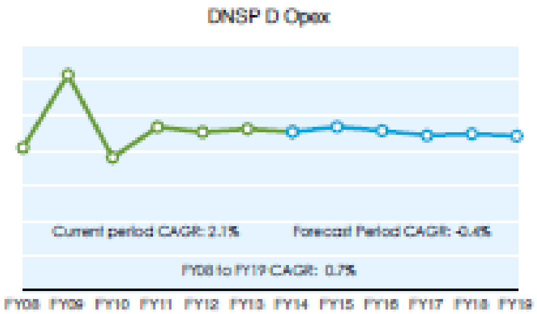
- Whether the circumstances impacting other businesses is driving trends relative to the DNSP. For instance, other DNSPs may be undertaking similar efficiencies at the same time, or may have found a need to increase their replacement programs.

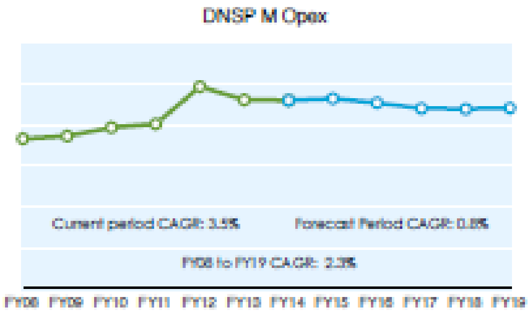
- We note that capex-opex substitution possibilities can impact the comparison over time. For instance, the reductions in capex may have an impact on opex forecasts.

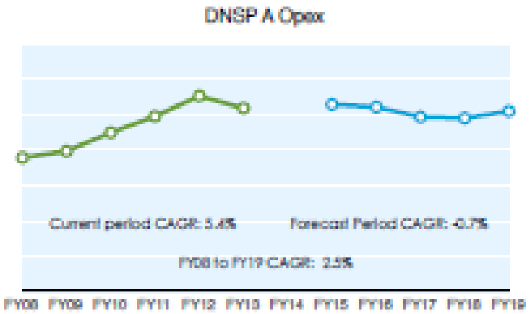
DNSP Opex Trends - Actual and Forecast

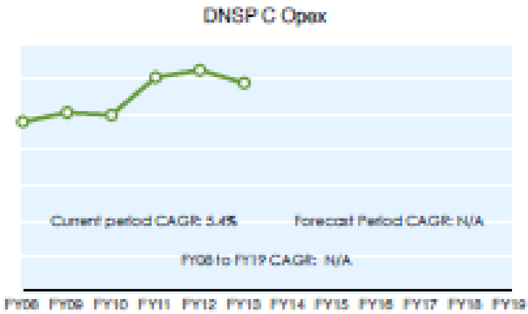


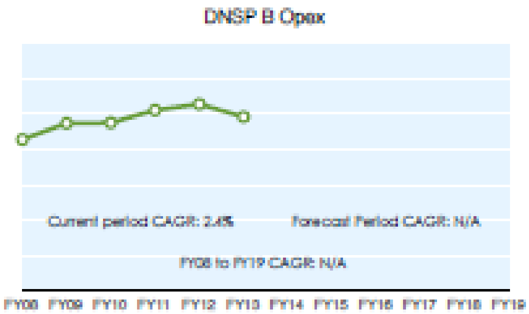












Notes:
1. All figures in FY2013 dollars.
2. Data sources are the RINs and directly supplied data from DNSPs.

4. Category level benchmarking

The AER notes that category level benchmarking allows it to compare expenditure across NSPs for categories at various levels of expenditure. Such benchmarking can provide granular cost assessments at the total level, or relative to operating conditions such as vegetation management costs per kilometer. Cost category analysis can also be used as a predictive model such as is the case for the AER's repex and augex models.

4.1 Effectiveness of technique in guiding decision making

In our view category analysis partially meets some of the criteria identified in the Productivity Commission's report, but even this is dependent on the manner in which the tool is being used. For the most part, we consider that the benchmark outcomes will be highly unreliable and cannot be used to infer relative inefficiency in isolation.

- Validity – Theoretically, cost category analysis can lead to more valid conclusions on efficiency of operations than high level benchmarking tools such as economic and aggregate analysis. This rests on the assumption that cost categories are reported consistently across DNSPs and that the data can take into account the drivers of expenditure. From a practical perspective these assumptions do not hold true:

    - The DNSPs in Australia have historically defined and recorded costs using different categorisations. The AER has sought to solve the issue of inconsistency by defining common categories in the RIN for benchmarking purposes. However, the definitions are still open to considerable interpretation, and the data provided will not be on a 'like for like' basis due to historical accounting practices. For example, overheads can be interpreted in many ways and will depend on the intrinsic way in which a DNSP has recorded costs in the past.

    - In addition to definitional issues, many of the cost categories are highly inter-related with each other such as maintenance and replacement expenditure. If the category is seen in isolation, it may mislead an analyst to conclude that a DNSP is inefficient in one area, and highly efficient in another.

    - Even within cost categories there is often limited financial information to adequately account for a particular driver of expenditure. For example, data on distribution network projects generally are recorded in bulk on financial systems, and therefore there is limited ability to identify whether drivers of investment relate to a particular issue with a technology type of a local driver such as a large new customer connection.

    - Finally there is no statistical method available to normalise for drivers underlying apparent cost differences. For instance, there is no sound method to account for the different costs that a DNSP incurs from constructing an asset in the CBD compared to an urban area.

- Accuracy and reliability – In many cases, the data can be verified with reference to financial systems and statements, and are accurate and reliable for that DNSP. However, due to the definition and interpretation issues identified above, the data cannot be relied on for the purposes of comparative data. Further, the AER's cost categorisations require DNSPs to map historical data and in some cases use 'best estimates'. Key examples are in splits by cost types of overheads which may draw on 'rough methods' due to insufficient data in financial systems. Given that this type of information is not drawn from actual data, there will be occasions when the data is not accurate or reliable to use for comparison.

- Robustness – As identified above, we consider that the data cannot be provided in a 'like for like' form and therefore cannot adequately control or normalise for operating differences between DNSPs.

- Manipulation of data – The AER's requirement that DNSPs provide best estimates where actual data cannot be provided, together with ambiguity on definitions may leave it open to a DNSP to report in a manner that puts its situation in the best light. This underscores issues with the reliability and accuracy of data when comparisons are drawn between businesses.

- Parsimony – To a degree, the AER's method of splitting opex and capex into cost categories is relatively straight forward. However, the AER has requested a high degree of granularity in data such as costs split by labour, materials and contractors. In our view this adds an additional degree of complexity that may lead the AER to form erroneous conclusions, and therefore do not meet the objective of parsimony.

- Fit for purpose – As explained below, cost category analysis may be fit for purpose, if used as a guide to the AER's detailed assessment of programs and projects rather than as a determinant or substitute for expenditure.

Despite not meeting the Productivity Commission's criteria, we consider that category benchmarking is potentially the best tool for a DNSP and regulator to apply when addressing the capex factor for benchmarking. This is because it allows the business and AER to consider whether there is a particular driver underlying cost differences (higher or lower) than peer DNSPs, or where the costs of the DNSPs are changing compared to trend levels for a well explained reason.

In these cases, further examination should be undertaken of the high level cost drivers that may explain the variance. For example, a DNSP's replacement costs may be explained by the relative age or failure rates of the assets on its network, or a new safety standard for a jurisdiction. If variances cannot be explained by different cost drivers, then this would signal that further work needs to be undertaken to assess the reasons for the variance. In this case, the AER may seek to identify the forecast method that was implemented by the DNSP, and the application of that forecasting approach to investment programs and projects. For example, an apparently high unit cost may be explained by a large project that is conducted in a major urban centre which provides a false picture of cost trends over time.

In the sections below we show how available benchmarks have informed the development and review of our forecasts. We provide information on cost category data, and also address the repex and augex models.

Endeavour Energy's forecasts have been heavily influenced by internal efficiency programs and industry reform that has focused on customer affordability. Comparative data between the 3 NSW DNSPs and our industry peers have played a role in identifying areas of efficiency, although this has been limited by the inherent issues with undertaking benchmarking. Our experience is that granular data can often paint a misleading picture on the relative efficiency in an area. Even when assessing data amongst the 3 NSW DNSPs, we noted that variances were impacted by definition and cost accounting issues. For example, even simple metrics such as travel costs which form a component of overheads, could not be normalised given that the underlying drivers across the 3 DNSPs are so different. For instance, a rural DNSP such as Essential Energy is likely to have higher relative transport costs per employee than Ausgrid.

For this reason, the reform process and our review of our forecasts have used comparative data with a high degree of caution. Where data has been assessed, we have not taken a simplistic view of assuming that the variance relates to efficiency. Rather we have undertaken a 'bottom up' assessment of underlying policies, forecast methods and cost controls of the DNSPs, in combination with available data and the ground experience of our staff. In this way, we were able to precisely identify the actions within the control of management to deliver efficiencies, whilst continuing to deliver services to our customers.

For the purposes of addressing the benchmarking criteria, we commissioned Huegin to provide cost category comparisons of the benchmarking group using 2009-14 and 2014-19 data if available. This follows from an earlier benchmarking study conducted by Huegin in 2012. Huegin's analysis is framed around understanding potential cost drivers and definitional issues underlying variances in costs between DNSPs. Further, we have examined the validity of specific cost category models such as augex and repex, including more detailed reviews of the model in Attachment 5.07 and 5.08 respectively.

4.2 Analysis of available data

Huegin Consulting undertook analysis of Endeavour Energy's expenditure levels for key categories relative to the 6 other DNSPs in the study. In the section below we set out the key results from the study for capex categories and opex categories.

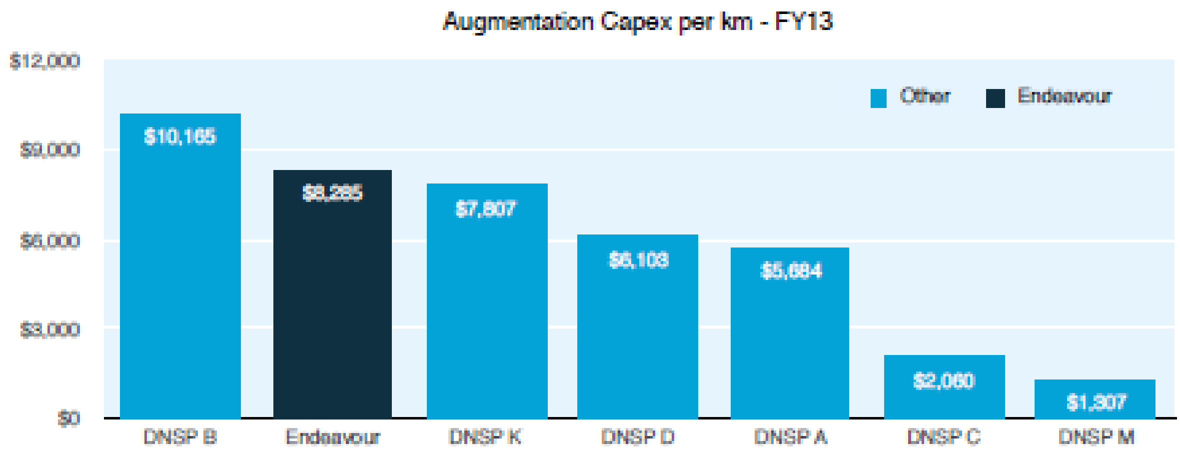
Capex

The Huegin study assesses the expenditure of DNSPs across 3 broad categories – augmentation, replacement and non-system capex. We note that these definitions are not common across DNSPs and in some cases a project may be driven by a combination of these drivers.

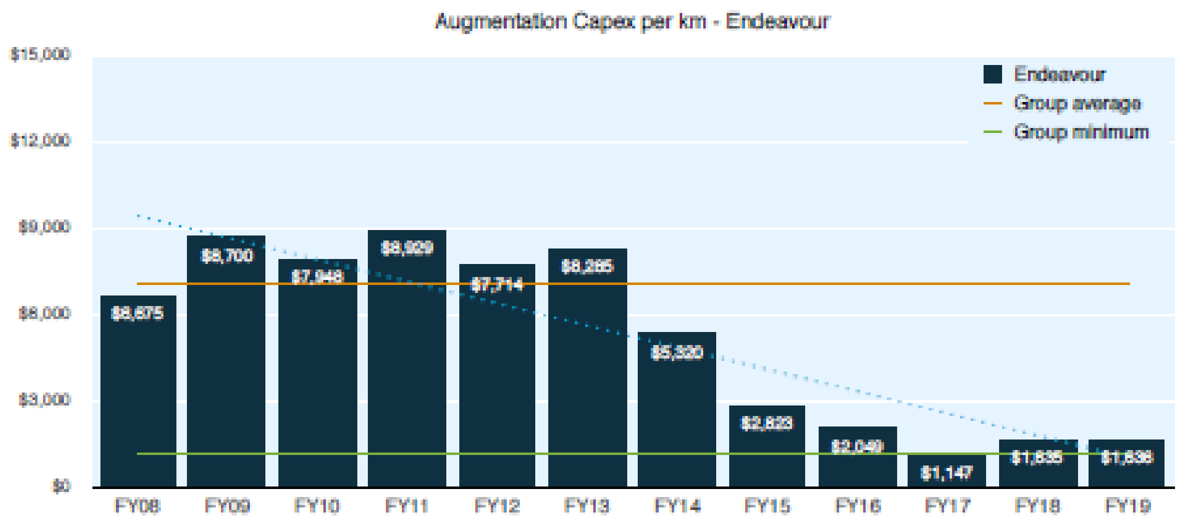
Augmentation expenditure

The AER note that this category of expenditure typically involves augmenting network components to ensure they have sufficient capacity to meet forecast demand. The AER have not been entirely clear about whether augmentation also includes works required to connect a new customer, as the AER refer to this as customer initiated connections. Huegin's report includes both customer initiated capex and reinforcements of the network for standard control services.

Huegin's study finds that Endeavour Energy's costs are the second highest in the group on a per kilometer basis. This relates more to the underlying drivers of augmentation at Endeavour Energy; for example the fact that Endeavour Energy's network area contains Sydney's North West and South West Growth Centres, widely regarded as the fastest growing corridor in the state.



Endeavour Energy's augmentation expenditure is forecast to be considerably lower in the 2014-19 period than in 2009-14. This reduction in spend is associated with the industry reform process, and that we had made significant inroads into meeting the security of supply criteria mandated in our licence conditions. They also reflect lower growth in peak demand compared to historical levels.



In addition to the analysis of augmentation expenditure, the AER has indicated that it will also be using the augex model as part of its cost category analysis. The augex model compares asset utilisation thresholds with forecasts of maximum demand to identify the parts of a network segment that may require augmentation. The model then uses capacity factors to calculate required augmentation, and unit costs to derive an augex forecast for the DNSP over a given period. The model is applied to segments of the network such as the sub-transmission network, 11kV network and low voltage network.

While we have sought to undertake preliminary analysis, we have noted that data limitations and uncertainty on the AER's preferred segmentation have limited its use. For this reason our response focuses on the effectiveness of the model from a conceptual point of view.

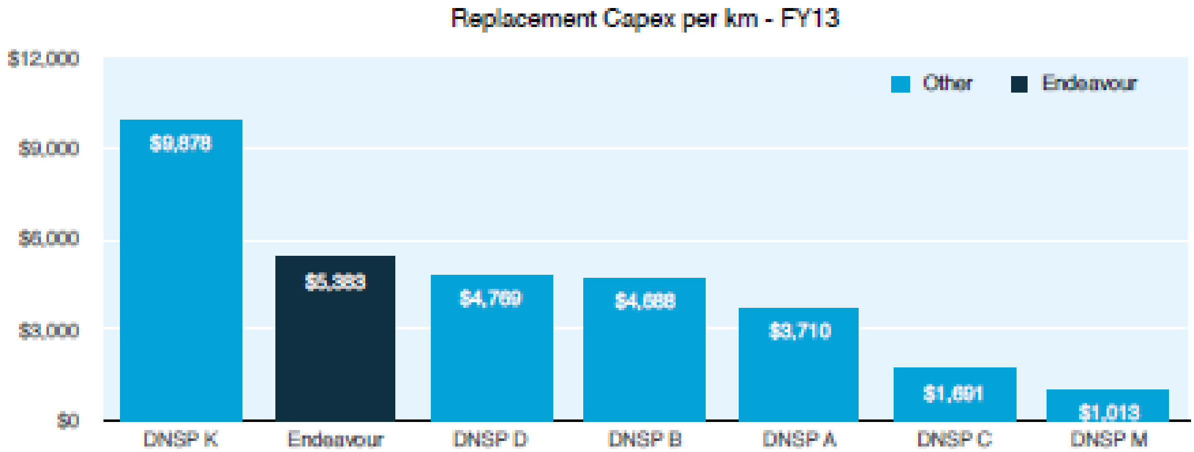
The NSW DNSPs have prepared joint analysis which assesses the effectiveness of the augex model as a benchmarking tool (please see Attachment 5.07). Our key finding is that the model cannot be used to develop an alternative or substitute forecast due to deficiencies in functional form and data limitations.

We note that the majority of our augmentation expenditure in the 2014-19 period is related to the need to provide infrastructure to service the greenfield development areas of North West and South West Sydney. The augex model specifically excludes expenditure on the connection of new customers and as such will be of limited usefulness to the AER in assessing our augmentation expenditure forecast. More information can be found in Attachment 5.07.

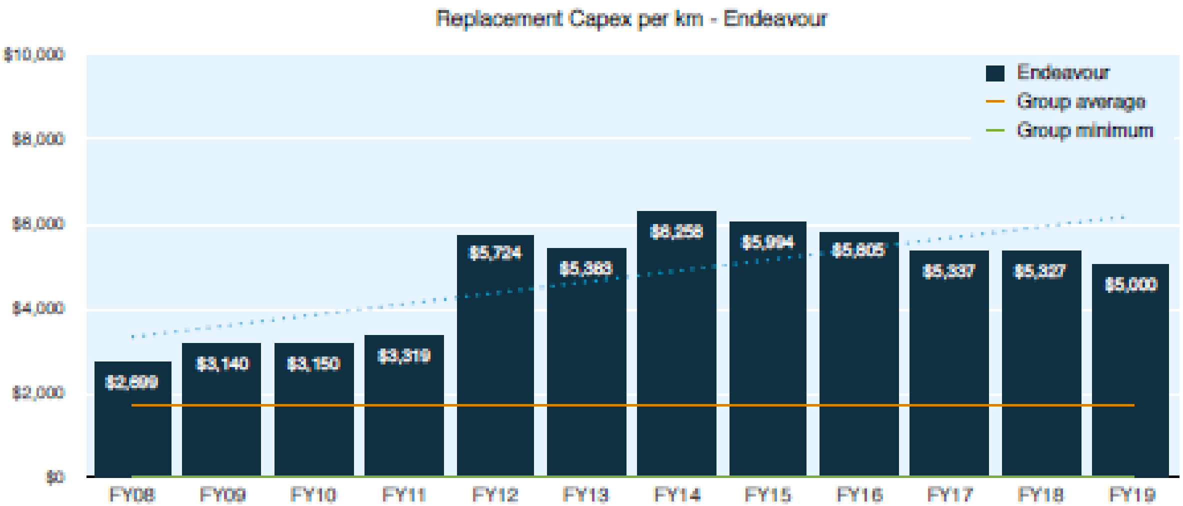
Replacement

The AER defines replacement expenditure as the non-demand driven capex to replace an asset with its modern equivalent where the asset has reached the end of its economic life. Economic life is determined by the age, condition, technology or environment of the existing asset. The capital expenditure is regarded as replacement expenditure if it is primarily determined by the existing assets ability to efficiently maintain its service performance requirement.

Huegin note that replacement capex for Endeavour Energy is within the range of other DNSPs on a per kilometer basis, however this also reflects the underlying characteristics of our network.



The graph below compares Endeavour's replacement capex in the 2014-19 period relative to the 2009-14 period. It shows that Endeavour is forecasting a higher level of capex relative to the current period, but is consistent with trends in the latter years of the period. Further, replacement capex starts to fall significantly in the latter years of the 2014-19 period.



In addition to replacement cost analysis, the AER will also be using the repex model. It has been described by the AER as a high-level probability based model that forecasts repex for various asset categories based on their condition (using age as a proxy) and unit costs. The AER has used it in determinations to compare NSP forecasts with the repex model outputs to identify and target areas in its forecast program that required detailed engineering and business case review.

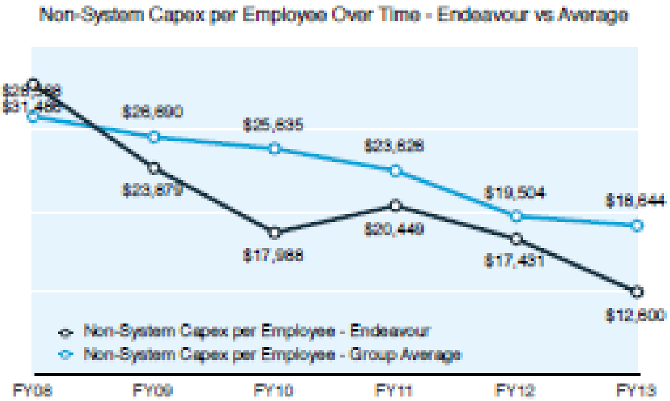
Endeavour Energy has, for a number of years, used the VDA (Value Development Algorithm) tool which, conceptually, is very similar to the repex tool, to guide the development of our replacement expenditure forecasts. We aim to maintain a sustainable level of replacement expenditure and use the VDA to provide us with a top down view of the appropriate level of expenditure to maintain the weighted average remaining life of the network at a consistent level. Bottom up expenditure forecasts are developed based on the known condition of specific assets and asset classes. The expenditure profile shown above is a function of the age profile of our network asset and the desire for a level of expenditure that will result in a sustainable network in the long term.

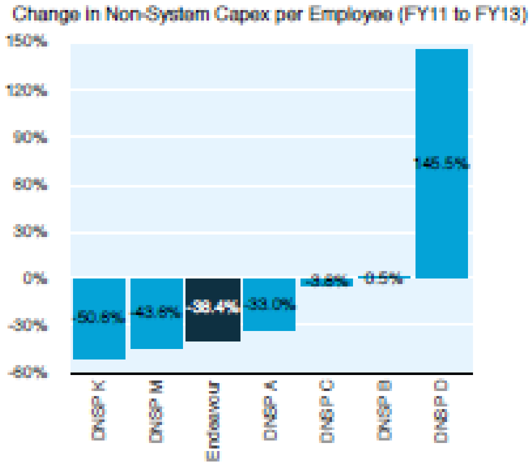
The NSW DNSPs have prepared joint analysis which assesses the effectiveness of the repex model as a benchmarking tool (please see Attachment 5.07). Our key finding is that the model cannot be used to develop an alternative or substitute forecast due to deficiencies in functional form and data limitations. Based on our experience with the VDA tool however, we note that the tool may have limited use as an informative tool for particular asset categories when there is sufficient population size, stability in replacement cycles over time, uniformity in technology type, and the costs are relatively stable over the population size.

Non system capex

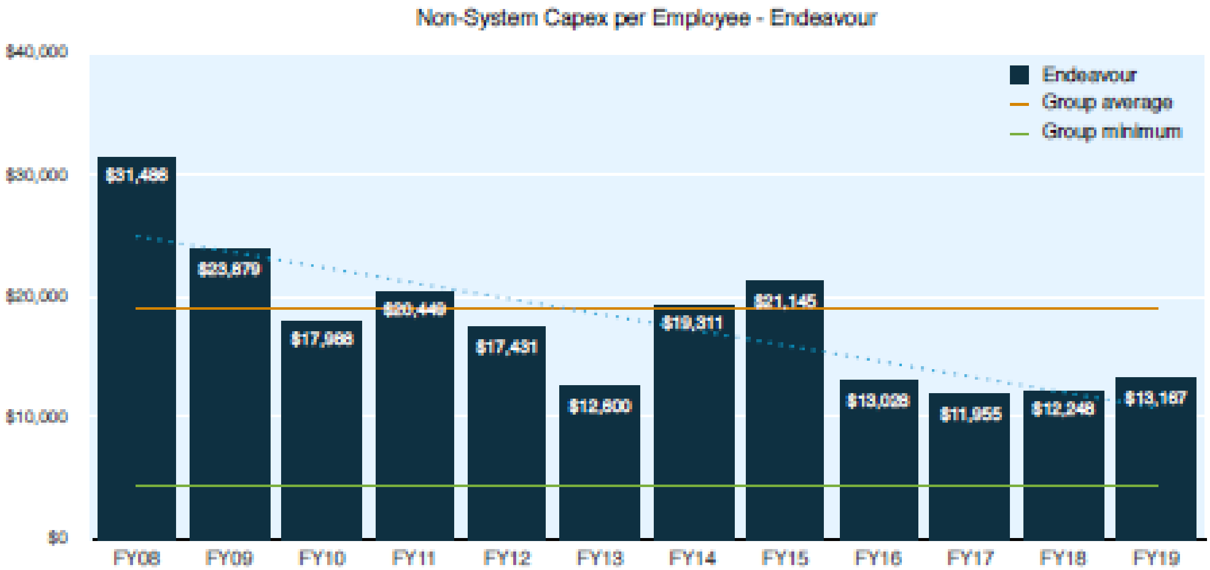
The AER identified the following types of non-network: IT, motor vehicles, property, SCADA and network control expenditure. Huegin have assessed relative levels of non-system capex based broadly on these categories of expenditure.

Endeavour Energy's non-system capex was significantly higher than other DNSPs in the study in the first 3 years of the 2014-19 period, but its expenditure levels have shown a significant decline since 2011-12. As can be seen in the graph below, this trend continues in the 2014-19 period, with average expenditure levels declining in comparison to the last 2 years of the regulatory period.





This indicates that efficiency programs and industry reform has delivered significant results in reducing non-system capex through changes in strategy and leveraging non-system capex across the 3 DNSPs. It also may be indicative of the high levels of capex in the early years of the period that delivered new functionality to support the larger capex programs, and the change in strategy which changes the focus of expenditure from delivering new capabilities to maintaining existing functions.



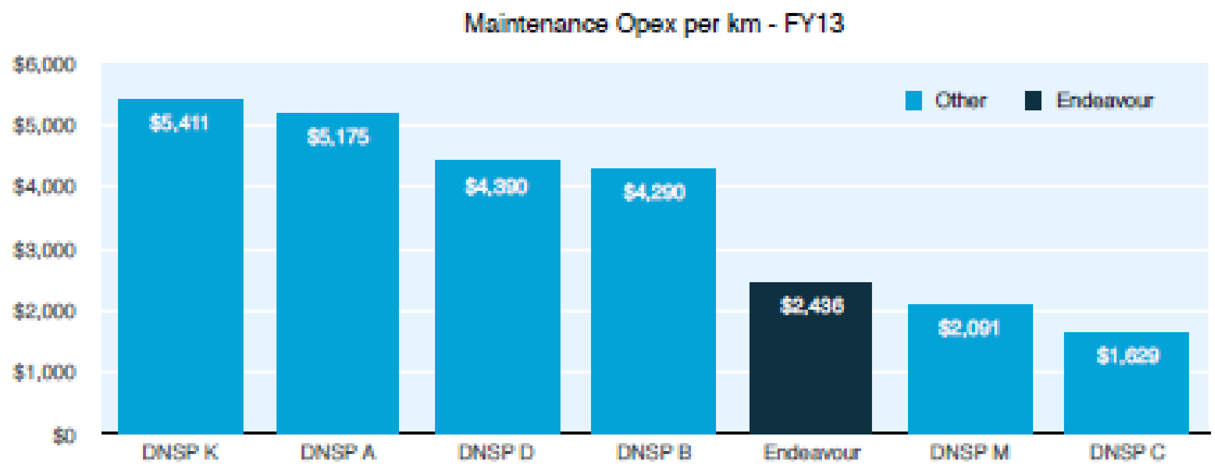
Opex

The Huegin study assesses opex of DNSPs across 2 broad categories – maintenance and operations costs. We consider that relative costs among DNSPs are likely to be shaped by capitalisation policies, age and other factors outside the control of management.

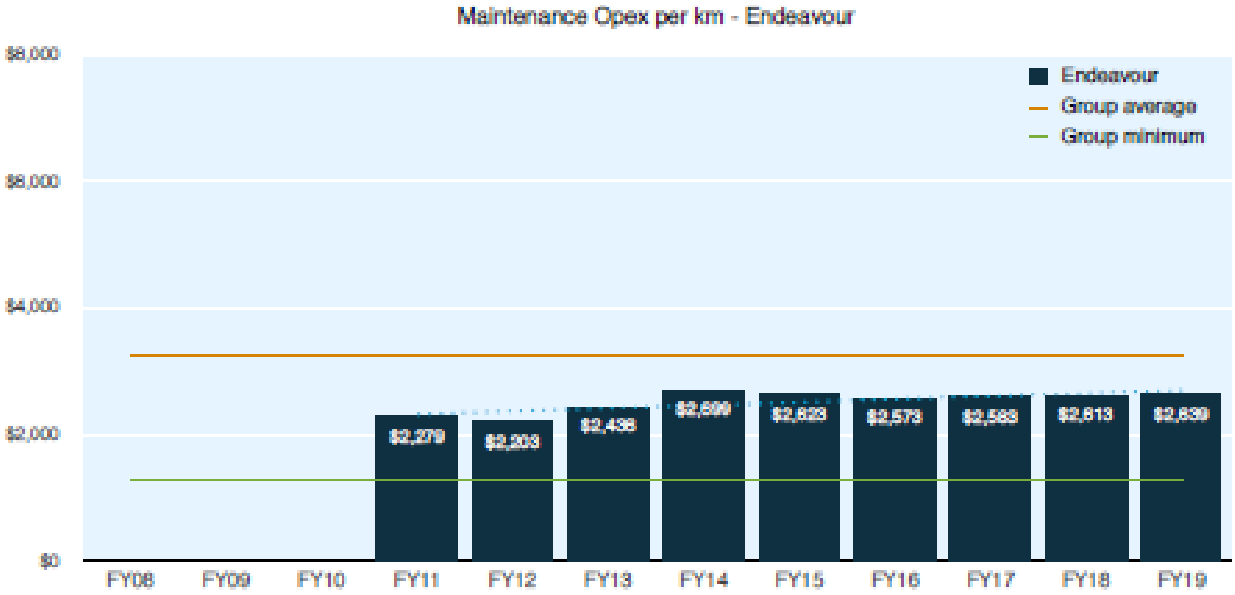
Maintenance

The AER define maintenance and emergency response to include all works to maintain the current working condition of an asset or to address the deterioration of an asset. These works include those that may be driven by reliability deterioration or an assessment of increasing risk of failure or performance degradation of a network asset. The AER consider that maintenance related to vegetation management is a separate category of expenditure defined as the process of keeping trees and other vegetation clear of electricity lines to reduce related outages and the potential for fire starts. Vegetation management also includes clearing easements and access tracks associated with electrical assets. Huegin's analysis includes vegetation management as part of its definition of maintenance.

Huegin's analysis shows that Endeavour Energy's maintenance costs on a per kilometer basis is among the lowest of the group in the study. Huegin note that there are a number of costs drivers of maintenance. For instance, long and radial networks in regional areas carry a significant cost premium in travelling between assets. At the same time, costs can be higher for DNSPs such as Ausgrid that operate in dense, urban areas where there is traffic congestion and greater sharing of common infrastructure.



Huegin also compared the trend in Endeavour Energy's performance from 2009-10 to 2014-19 and compared this to forecasts for the 2014-19 period. The diagram shows that Endeavour Energy's maintenance expenditure is below the group average, and is relatively flat over time.

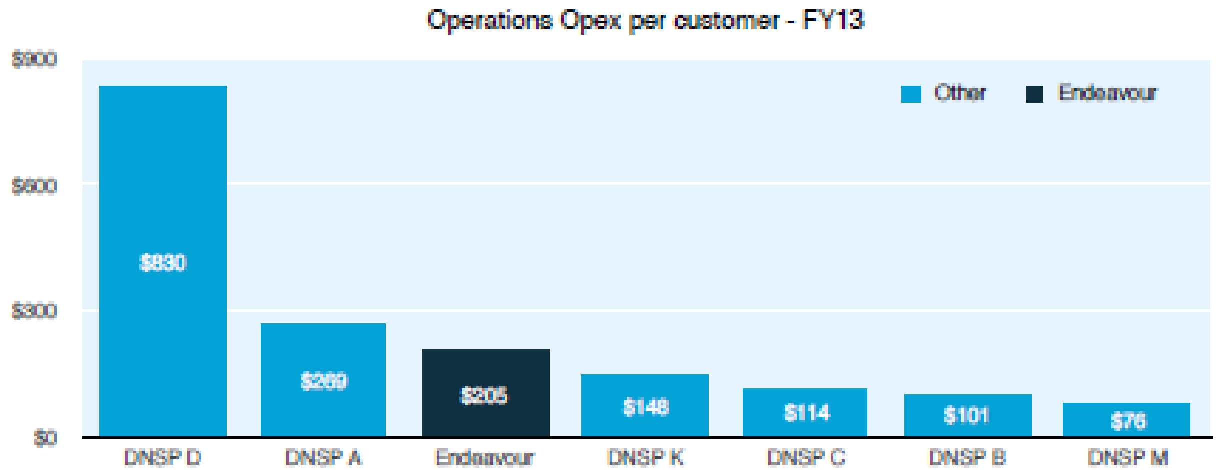


Operations opex

Huegin define operations opex as network control, systems operations, customer operations and support functions such as IT, property and fleet management. This broadly relates to the AER's definition of non-network costs.

Huegin's operations costs for this study have been compared using the number of customers as the comparison basis. Comparisons of operations costs per customer for the 2013 financial year are shown below.

Huegin's analysis shows that Endeavour Energy's operations opex is the second lowest of the DNSPs surveyed, Huegin note that operating costs are largely driven by the location and complexity of the network, customer base and the business scale. This makes comparison particularly difficult, as the number and relative influence of cost drivers varies across businesses.



Huegin also compared the trend in Endeavour Energy's performance from 2009-10 to 2014-19 and compared this to forecasts for the 2014-19 period. The diagram shows that Endeavour Energy's operating costs are below the group average in the 2009-14 period and forecast to steadily decline further below the average during the 2014-19 period. We consider there are a range of drivers influencing lower costs including efficiency programs and industry reform.

Operations Opex per Customer - Endeavour