

[C-i-C begins] = information not to be released without a confidentiality undertaking
[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

Report on:

Economic Insights ‘Testing Further Specifications’ for the 2015 DTCS FAD

by

Professor Trevor Breusch

02 6629 6042

trevor.breusch@gmail.com

8 February 2016

1. Introduction

Background

- 1.1 The Australian Competition and Consumer Commission (“ACCC”) has provided me with another report by Economic Insights (“EI”), this time entitled *Domestic Transmission Capacity Services Benchmarking Model: Testing Further Specifications: Report prepared for Australian Competition and Consumer Commission* and dated 16 December 2015 (“**EI Further report**”). A revised version of the EI Further report with a date of 18 January 2016 was also provided.
- 1.2 The EI Further report is the latest in a series of technical reports on benchmarking the DTCS, including a workshop discussion paper dated April 2015, a draft report on the benchmark modelling work dated June 2015, and a final report dated 1 September 2015 (“**EI Final report**”). I have provided reports to King & Wood Mallesons on behalf of Telstra in relation to each of these reports by EI. My report on the EI Final report (“**Breusch on Final**”) also included my comments on the ACCC’s Draft Final Access Determination (“**Draft FAD**”). I understand that all of my earlier reports have been submitted by Telstra as part of its responses to the ACCC’s requests for feedback from stakeholders.
- 1.3 The ACCC has given me access to various files of data, the most recent of which is entitled *2015 DTCS FAD - revised data set including Optus VHA eJV joint venture pricing - commercial in confidence CIC v6 - GST de-identified for experts.xlsx* (“**ACCC data file**”), which I downloaded on 23 December 2015. I have also been given hundreds of other files containing EI’s computer code, transformed data and intermediate results as used to prepare the formal reports. I have been able to replicate most of the econometric research undertaken by EI and reported in its Workshop, Draft, Final and Further reports.

Purpose of this report

- 1.4 This report responds to a request in a letter dated 1 May 2015 from King & Wood Mallesons for me to provide expert reports on the benchmarking framework proposed by EI and adopted by the ACCC for the 2015 FAD.
- 1.5 In addressing the very large volume of material contained in EI's four reports (and revisions), I will focus attention on the EI Further report while taking the earlier reports largely as given. As stated in paragraph 1.2 above, I have already commented in detail on these earlier EI reports. However, there are a few outstanding issues from the EI Final report and the Draft FAD which I will comment upon in this report, to ensure they are not overlooked in the implementation of benchmarking in the eventual FAD.
- 1.6 This report is structured as follows:
 - (a) **Section 2:** Summary of my findings in relation to the EI Further report
 - (b) **Section 3:** Mistakes in the EI Further report
 - (c) **Section 4:** Expanding the data set to include the Joint Venture data
 - (d) **Section 5:** Outliers (EI Further report, Chapter 3)
 - (e) **Section 6:** 2Mbps services (EI Further report, Chapter 4)
 - (f) **Section 7:** Stochastic frontier analysis (EI Further report, Chapter 5)
 - (g) **Section 8:** Issues carrying over from the EI Final report/Draft FAD
 - (h) **Section 9:** Overview of benchmarking the 2015 DTCS FAD
 - (i) **Section 10:** Conclusions
 - (j) **Section 11:** References.

2. Summary of my findings in relation to the EI Further report

- 2.1 I begin by restating the objective in benchmarking the DTCS. A benchmarking model is required to represent the overall or average relationship between the charge for a service on the exempt (that is, competitive) routes and the observable characteristics of that service. The regulated (maximum) charge on a declared route is obtained by predicting what the charge would be on the counterfactual assumption that it is an exempt route with the same relevant characteristics. Various simplifications may be applied to the statistical prediction to obtain a workable pricing formula.
- 2.2 I had observed that the EI Final report made good progress towards a benchmarking model for price regulation in the DTCS, despite the sometimes arduous path of the model developments in earlier EI reports. My principal criticisms of the earlier reports related to the use of inappropriate statistical criteria, which were more suited to statistical

[C-i-C begins] = information not to be released without a confidentiality undertaking
[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

inference in academic research than finding a benchmarking model, and to unsafe use of context ‘knowledge’, which was often of dubious validity and consisted of no more than ex-post rationalisations of statistical results. I observed that the models of the EI Final report are much simpler than the models entertained in earlier reports from EI, containing fewer variables of questionable economic and statistical relevance. As a result, they are more easily understood and translate more readily to benchmarking formulas. I proposed that the models of the EI Final report could, with minor adaptation and correction, be used for benchmarking the DTCS.

- 2.3 The EI Further report examines a range of issues indicated as concerns by stakeholder responses to the earlier EI reports. The principal such issues are: possible outliers in the data; the suggestion that low-capacity, short-distance services are priced too high; and a proposal to use stochastic frontier analysis instead of regression analysis to form a benchmark pricing equation. I will consider each of these issues in detail in separate sections of this report. I endorse most of EI’s findings in the EI Further report and, perhaps more strongly than EI, I conclude that the preferred models of the EI Final report dominate any of these alternatives. In that regard, the EI Further report can be seen as a waste of time and resources.

Complexities in modelling and lack of transparency

- 2.4 Despite largely sharing EI’s overall conclusions, I find there are a number of errors in the EI Further report, particularly in the form of discrepancies between the method stated in the text and the method that is actually employed to achieve the results reported in tables. Some of the detailed comparisons drawn by EI are erroneous and misleading as a result. The importance of identifying these mistakes is not that it changes my view that EI’s conclusions are largely correct, but rather that it shows how the modelling has become so complicated that even the skilled data analysts in EI and the ACCC cannot see the implications of their calculations. If such skilled data analysts do not understand what is being done to produce regulated prices, there is very little prospect that other stakeholders can consider or respond fully to key inputs into the FAD, or have confidence in the outcome. Such complexity and the resulting lack of transparency have no place in public policy formulation.

Joint venture data

- 2.5 The latest version of the data set used for modelling includes an additional 941 observations that were provided late in the process. These observations come from a joint venture (JV) in which Optus is the supplier and VHA is the client. In Breusch on Final, I speculated on possible difficulties that might result from including this additional data, and now that the missing variables relating to ESA and route characteristics have been included in the data file, I can comment more fully. There is strong and clear

[C-i-C begins] = information not to be released without a confidentiality undertaking

[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

evidence that these additional observations do not come from the same population, and do not have the same relationship to the driver variables, as the earlier data. Apart from the [C-i-C starts] [Confidential] [C-i-C ends], there is extensive price averaging between routes with very different characteristics. As a result, the prices in this large bundle will give a distorted picture of the economic supply price of a service with a specified set of characteristics. Aside from any issues of procedural fairness in adding this additional data so late in the process, and real doubts about the competitive nature of prices in a joint venture arrangement, I conclude that benchmarking the DTCS should exclude this JV data because it is so different from the other data collected for the purpose by the ACCC.

Choice of model

2.6 The EI Final report provides two models and their associated pricing formulas, called Model 2 and Model 3 (also called respectively Models A1b and A1c in the EI Further report). The ACCC adopts Model 2 as the basis of the Draft FAD. Contrary to the ACCC's choice, I found in Breusch on Final that economic logic, statistical adequacy and transparency of process all indicate that Model 3 (or A1c) is the better choice. Nothing in the EI Further report causes me to vary from that opinion; indeed there is some additional evidence that my earlier preference for Model 3 (or A1c) is correct.

Other issues from the EI Final report

2.7 Some issues from the EI Final report have not been addressed in the EI Further report, but should not be forgotten. These are: the 'contract start date' variable is of doubtful quality and should not be used constructively (which excludes the use of Model 1); 'ESA throughput' variable is mostly irrelevant to DTCS and has a perverse effect in the models (which argues against both Models 1 and 2); the assumed distance setting of 2km for tail-ends is both arbitrary and its justification is lacking in mathematical logic; the connection charges pricing model of the Draft FAD is deficient; and the ACCC's proposal to price SDH services as if they are Ethernet is lacking in economic rationale. These are all noted in more detail below.

2.8 As of late September 2015, the process of benchmarking the DTCS appeared to be converging on an outcome described in the EI Final report and the Draft FAD. The EI Further report has opened out this process in ways that may be hard to bring to a conclusion. There are now many more methods and models on the table, with seven new statistical specifications (Models 1-7), each with three different forms of the predictor variables (variants a, b and c), and each combination being estimated with and without the additional joint venture data. Moreover, each of these combinations is explored in the EI Further report for its pricing implications in detailed segments of the DTCS market. There is a risk this large menu of alternatives will aggravate the extent of stakeholder

[C-i-C begins] = information not to be released without a confidentiality undertaking

[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

disagreement and lobbying, including “cherry picking” of methods based on seeing favourable results in particular market segments.

3. Mistakes in the EI Further report

- 3.1 There are reporting discrepancies and statistical errors in the EI Further report, which render some of the model comparisons and recommendations invalid. Three levels of mistake can be found: one is where an assumption is employed but that assumption is not documented anywhere accessible to the reader of the report; a second kind of mistake is where some adjustments are made partially but incompletely, with no apparent logic to where the boundary is drawn; and the third kind is erroneous statistical method. Sometimes the explanation in the text is so vague that I am uncertain which kind of mistake might be involved in each particular instance.
- 3.2 All of the discrepancies described in this section concern the prediction of average prices for declared services in various segments of the market, where those segments are distinguished by the various route types (inter-capital, metro, regional and tail-end). I found these discrepancies by replicating the estimates of the various models and then attempting to replicate the predicted prices as they are tabulated. I have only done this experiment for Models 1, 4 and 5, so the results to which I refer are in Tables 2.5, 4.5 and 4.11, respectively. However, it is clear that at least some of the same discrepancies will occur in other models.
- 3.3 Readers are advised “It is important to note that the models used for prediction are those simplified according to the methods described in our final report (Economic Insights 2015).” (EI Further report, p.74). However, this advice is of little use in the case of model variant ‘a’ (i.e. Models 1a, etc.). The equivalent Model 1 of the EI Final report is not a preferred model there, and hence it is not resolved to a formula for predicting prices. However, subsection 6.2.3 of the EI Final report does describe an example of how the ‘contract start date’ variable might be used constructively for a start date of 1 April 2016. My replication experiments indicate the tabulated prediction results require the start date to be set to 1 January 2016. I can find nothing in the EI Further report to inform the reader that is the date employed in their tables, although a close inspection of some of the many files of EI’s intermediate results sent to me by the ACCC has confirmed that is indeed the assumption. It seems this particular discrepancy is more an unstated assumption than an outright error in the calculation of predicted prices. However, the well-known problems of data quality with the ‘contract start date’ variable argues against using this variable to manipulate prices in this way. This separate matter is discussed in section 8 below.

[C-i-C begins] = information not to be released without a confidentiality undertaking
[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

Tail-end services

3.4 The method for predicting prices for tail-end services in the tables of the EI Further report also differs from the method described in the EI Final report (sections 6.2 and 6.3). Rather than attributing to each tail-end service the route effect of either the metro and regional ESA to which that tail-end belongs, the spreadsheets behind EI's tables of predicted prices calculate the route effect of a tail-end as an average across both ESA types. This method makes the elementary statistical error of calculating a simple unweighted average of the two route coefficients, despite there being nearly three times as many of the lower-value metro types than the higher-value regional types. As a result, the predicted prices for standalone tail-ends as shown in Table 2.5 for Model 1 are higher than would be obtained by the method described in the EI Final report. The same error occurs in all of the models I have replicated (all variants of Models 1, 4 and 5). A quick inspection of EI's files of intermediate results suggests it is universal across all of the predicted prices for tail-end services in the EI Further report. This mistake is both a failure to declare properly the method used *and* an error of statistical method.

Prices for 2Mbps services

- 3.5 An additional problem is found in the predicted prices from Model 4. In this class of models, the low-capacity/short-distance services (called "2Mbps" services) are omitted from the data set on which the regression model is estimated. Prices for the non-2Mbps services are predicted from the model in the same way as for Model 1. A separate prediction formula based on an average of the 2Mbps services on exempt routes is employed for pricing the 2Mbps services. (The actual method involves averaging the logarithm of the monthly charge and using that together with its standard deviation to construct an estimate of the average charge in dollars. That round-about method yields \$321.23 where the simple average monthly charge is \$323.23, so the specified method seems needlessly elaborate.)
- 3.6 The frame for the exempt 2Mbps services to be averaged is described as "short metro '2 Mbps' services" (p.38), although it is clear from my replications that *all* exempt 2Mbps services are used, whether metro or regional. (There are no inter-capital services in the 2Mbps category and no exempt tail-end services.) The frame for the declared services to be predicted is described as "short, low capacity, metro services" (p.39), but again the prediction is also applied to 2Mbps services on declared regional routes. So far, this discrepancy fits into the category I described where the actual prediction method does not match the description given in the text.
- 3.7 In a further discrepancy with Model 4 that is surely a *logical* error, the 2Mbps *tail-end* services in Table 4.5 are predicted from the regression model established for predicting the non-2Mbps services, not from the average of the exempt 2Mbps services. That

[C-i-C begins] = information not to be released without a confidentiality undertaking
[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

implies that the 2Mbps tail-ends are predicted at an average price of \$487, where the price for each 2Mbps non-tail-end on a metro or regional route is restricted to \$321. That arrangement is difficult to understand, and certainly not documented in the report. Coupled with the separate error I describe in paragraph 3.4 above, this mistake explains why the average predicted prices for tail-end services in Table 4.5 are so surprisingly high.

- 3.8 Model 5 includes a dummy variable for 2Mbps services, instead of restricting applicability of the regression model to the non-2Mbps cases and predicting the 2Mbps cases separately, as used in Model 4. Nevertheless, the predicted prices for tail-end services from Model 5 in Table 4.11 contain errors similar to those identified with Model 4. When Model 5 is used in prediction, the dummy variable that indicates 2Mbps services is only applied to 2Mbps *metro* and *regional* services. The 2Mbps *tail-end* services are predicted with the dummy variable suppressed, as if they were non-2Mbps services. Again this unexpected arrangement is not documented, and the result is predicted prices for tail-ends in Table 4.11 that are surprisingly high.
- 3.9 There may be a justification for predicting 2Mbps tail-end prices from Models 4 and 5 as if they are non-2Mbps services, but that reason is not given and I cannot guess it. Even if it is meaningful to predict tail-end prices this way, the predicted prices for tail-ends from Models 4 and 5 suffer the same statistical error described for Model 1 in paragraph 3.4 above. Other models in the EI Further report may have their own peculiarities in prediction similar to those of Models 4 and 5, but I have not discovered them because I have not attempted to replicate EI's calculations of the other models.

Conclusions on modelling complexity

- 3.10 The implication of this discussion of discrepancies between the methods documented and the methods actually used is *not* that some minor repairs are needed. The important lesson is that the modelling has become so complicated that EI and the ACCC have lost sight of the models and their implications for benchmarking. If such skilled data analysts do not understand what is being done to produce regulated prices, there is very little prospect that other stakeholders – let alone the public at large – will have confidence in the outcome. Public policy demands a simpler, more transparent, process.

4. Expanding the data set to include the Joint Venture data

- 4.1 The data set used in the EI Further report contains an additional 2015 observations, which became available around the time the EI Final report was issued. These observations come from a joint venture (JV) arrangement in which Optus is the supplier and VHA is the client for these services. In response to a request from the ACCC, I commented briefly in Breusch on Final regarding the appropriateness of using this

[C-i-C begins] = information not to be released without a confidentiality undertaking

[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

additional data for benchmarking. At that time the exchange service areas (ESAs) in the data file were not identified, hence none of the variables derived from ESAs (route type, route and ESA throughput, declared/exempt indicator, distance, etc.) were included in the file. The data set now includes these variables, so there are 941 additional observations on exempt routes that might be used to specify and estimate a benchmarking regression model.

4.2 I noted in Breusch on Final that the additional observations have two characteristics that differ from the original data. [C-i-C begins]

[C-i-C ends]

[C-i-C ends] Since these are key driver variables in the candidate benchmarking models of the EI Final report, I speculated that the additional observations would not be consistent with the models of EI Final report, and that the modelling would have a very different outcome if the additional data were included.

4.3 Now that the variables that depend on ESA identification are available, it is possible to revisit my earlier speculation and test some of the issues I raised. The simplest such investigation of a possible structural break is to estimate the model on the full exempt data set including the 941 additional JV observations, but to include a dummy variable to indicate the JV observations. The results of this test vary slightly with the specification of the model, but taking Model 1b as the central illustration, the outcome is a coefficient on the dummy variable of [C-i-C begins] [Confidential] [C-i-C ends] and an associated [C-i-C begins] [Confidential] [C-i-C ends]. That indicates a price drop of [C-i-C begins] [Confidential] [C-i-C ends] per cent when measured logarithmically, which is equivalent to a [C-i-C begins] [Confidential] [C-i-C ends] per cent drop when measured as a proportion of the higher price level.¹ That price drop is strongly statistically significant, well beyond any conventional level of significance. Indeed, this structural break is more strongly statistically significant than many of the variables retained in Model 1b for their statistical relevance, such as the logarithms of route throughput and ESA throughput, or the indicator for SDH interface technology.

4.4 Another standard test for a structural break between the original and JV data consists of estimating in the two subsamples separately and comparing their joint results with estimates using the pooled data. In this test both the overall level and all of the coefficients are free to change, whereas in paragraph 4.3 the coefficients were held constant throughout and only the overall level was used to represent the difference between the two subsamples of data. Some of the required material for this second test is provided in the EI Further report: for instance Model A1b is estimated on the original

¹ It is standard econometric practice to measure proportional changes as the difference in logarithms. This method is used extensively throughout the series of EI reports. It is explained in some detail in a text box in my response to the EI Workshop report.

[C-i-C begins] = information not to be released without a confidentiality undertaking
[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

subsample excluding the additional data but otherwise follows the specification of Model 1b estimated on the pooled data. The same model structure can be estimated using only the additional JV observations, and a likelihood-ratio test performed. The outcome again is an overwhelming rejection of the hypothesis that the data come from the same population, with statistical significance well beyond any conventional level.

- 4.5 It is instructive to investigate the model that is estimated using only the additional 941 observations of JV data on exempt routes (as required for the likelihood-ratio test in the previous paragraph). Using Model 1b again to determine the model structure, [C-i-C begins] [C-i-C ends] None of the other route or ESA characteristics, including expected key drivers such as [C-i-C begins] [C-i-C ends]. (It is obvious that the provider-specific indicators will play no role in these estimates, because all the data comes from the one supplier.) It is likely that the marginal significance of the logarithm of route throughput is simply a chance variation, so these results confirm the speculations I made in Breusch on Final.
- 4.6 The very large price discount of [C-i-C begins] [Confidential] [C-i-C ends] per cent against the reference provider observed in the joint venture data cannot be explained as a provider effect in the manner of the effects reported in the various models of EI Final and Further reports. Indeed, in the version of Model 1b with the dummy variable indicator for the JV data, the provider of JV services is seen to [C-i-C ends] [Confidential] [C-i-C ends] relative to the reference supplier in the original sample of data. This result confirms that same supplier has a very different price profile in the JV data from that observed in the competitive DTCS data.
- 4.7 A reasonable question is: Why does EI not find these clear indications that the additional data is so different from the original data in both overall price level and in the relationship of price to the key driver variables? The answer is that they do not look for, nor report on, evidence that the additional data comes from a different population. There are no separate estimates using just the additional data nor any tests of homogeneity of the two subsamples. However, there are reported estimates with and without the additional data, and hints of the different data sources can be seen by comparing estimation results. The effects of both of the key drivers, capacity and distance, on prices are attenuated when the JV data is included, as can be seen by comparing the “cost elasticities” at the 50th percentile, shown in Tables 2.8 and A.7. Similarly the coefficients of both of the route types, metro and regional, are smaller when the JV data are included. The formal test for a structural break with the additive dummy variable (as in paragraph 4.3 above), and the broader likelihood-ratio test (as in paragraph 4.4 above) should have been provided in the EI Further report. Doing so would clearly indicate the

[C-i-C begins] = information not to be released without a confidentiality undertaking
[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

invalidity of including the additional data when seeking to benchmark the DTCS to competitive prices, which, as set out in paragraph 2.1 above, should be the central objective of any benchmarking exercise.

- 4.8 Further understanding of the impact of including the JV data can be obtained by comparing predicted prices between models excluding or including the JV data in estimation, provided that comparison is made *for the same set of services*. That comparison is not possible between Tables A.4 and 2.5 of EI Further report, where the services used for prediction vary to correspond with the services used for estimation. Again taking Model 1b as the central case, the average predicted price for all 6,767 exempt services not in the JV data falls from \$1,207 in Table A.4 to \$1,110 (8.4 per cent lower) while for all 11,480 declared services not in the JV data it falls from \$834 to \$762 (9.0 per cent lower). Using Model 1c, the average predicted price for the exempt services falls from \$1,224 in Table A.4 to \$1,134 (7.6 per cent lower) while for the declared services it falls from \$865 to \$807 (6.9 per cent lower).²

5. Outliers (EI Further report, Chapter 3)

- 5.1 The EI Further report explores the matter of outliers, both as observations that might be removed from the data set when the statistical model is fitted and as observations that might be downweighted by choosing a “robust” estimation method. I largely agree with EI’s conclusions in regard to the identification and treatment of outliers, although I disagree with the criterion they have adopted that the treatment of outliers should reduce prices on low-capacity metro routes. (EI Further report, p.22)
- 5.2 I agree with EI that an observation should be removed as an outlier only if it does not contain relevant information or it contains measurement error. The data (at least in the original sample used in the EI Final report) was obtained by the ACCC against a specification of relevance and confirmed by the provider when doubts arose as to the validity of any observation that appeared mistaken.
- 5.3 Mechanical removal of outliers according to a criterion such Cook’s Distance (p.17) is problematic even in situations where deleting observations might be a reasonable strategy. If the data is contaminated, the first round estimates will not be accurate, so the initial decision to remove observations will be made on a misspecified model. There is no guarantee that iterating the procedure (i.e. updating the model and re-determining which observations are outliers) will converge. Even if the process is convergent there is

² These predicted prices on declared routes use the formula for pricing tail-end services as specified in the EI Final report, not the incorrect methods used in Table A.4 and the other tables of predicted prices in the EI Further report. The details of these errors are described in Section 3 above.

[C-i-C begins] = information not to be released without a confidentiality undertaking

[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

no guarantee that only the contaminated cases will be identified as the outliers. Further, at each stage a decision rule is needed to identify the likely outliers, and while there are suggested criteria in the literature, there is no correct way to do that. The outcome will depend on the operator's chosen criteria.

- 5.4 The received statistical literature on outliers appears less relevant when the objective is benchmarking rather than statistical inference. The standard treatment assumes that valid data from the target population is well described by a statistical distribution but the sample is contaminated by a small percentage of outlier values from some other source. That is not a convincing approach when the data are deliberately collected as information on which to form benchmark. Even when some observations are statistical outliers relative to the idealised statistical model, those observations would still contain valid information for benchmarking.
- 5.5 The alternative treatment of outliers by use of some "robust" regression procedure also finds me largely in agreement with EI. There are many methods available in the literature and, as EI note with clarity, the differences between them are idiosyncratic and the choice of method is largely subjective. It seems particularly unsafe to adopt a method derived from a recent PhD thesis and implemented in free public domain software that is expressly unwarranted. Again, several parameters typically need to be specified by the operator, thus allowing for variable and somewhat subjective results. These methods may have valid application in academic research, but they are particularly unsuited to the formation of transparent public policy.
- 5.6 I have chosen not to attempt replication of Models 2 and 3, because these models are clearly unsuited to the purpose of benchmarking the DTCS, irrespective of their estimation results or the predicted prices in the various market segments. Indeed, the tabulated price comparisons with other models and the associated discussion in the EI Further report seem more clearly directed at addressing Optus's complaint about the benchmark prices for low-capacity metro routes than any problem of substance relating to outliers in the data. This particular "problem" with the benchmark prices is discussed in section 6 below.

6. 2Mbps services (EI Further report, Chapter 4)

- 6.1 This chapter in the EI Further report pursues the suggestion by Optus that benchmark prices on low-capacity short-distance services are too high. As I noted in section 3 above when discussing some discrepancies and mistakes in the EI Further report, for Models 4 and 5 the particular services of interest are defined as having capacity less than 2.5 Mbps and distance less than 5km. For Model 6, it appears that the former capacity criterion is used but for technical reasons the distance criterion is ignored. In discussion in the text

[C-i-C begins] = information not to be released without a confidentiality undertaking

[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

of the EI Further report, the focus is often described as “2 Mbps metro” services or just “2Mbps” services.

- 6.2 The added dummy variable in Model 5 has quite a large coefficient and t-ratio (–.227 with $t = -9.31$ in Model 5b, as the middle case), which lends some support to the proposition that 2Mbps services on exempt routes are on average priced somewhat lower than the model would suggest. Similarly, excluding the 2Mbps services from the fitted model and using the sample average for them as a simple price predictor, as in Model 4, does yield lower price predictions for 2Mbps services. However, as EI observe, while treating these services differently might predict lower prices for them, it does so at the expense of higher prices for other services.
- 6.3 There are several dangers in the approaches of Chapter 4, which if allowed to persist will undermine the fundamental impartiality and transparency of benchmarking. It seems that EI has invented a new criterion in model selection – that the model should accurately predict prices on declared routes in a market segment of particular interest to a stakeholder. The isolation of one category of service (low-capacity, short-distance, metro) in which the stakeholder has a special concern will invite other stakeholders to indicate particular segments of the market where special treatment might be more favourable to them. Benchmarking models formulated by independent consultants for the ACCC should be immune to such “cherry-picking”.
- 6.4 There is nothing unusual in the finding that the 2Mbps services deviate from the fit of the model. If the same approach of adding an indicator variable to Model 1 (as done in Model 5) is taken with other groups of services, then similarly large and statistically significant deviations will be found. I will use Model 1b for illustration, not as an endorsement of that model but as the middle case of EI’s three model variants. With 200Mbps services indicated, the coefficient is –0.199 with $t = -7.75$, showing that 200Mbps services are also priced lower than the model would suggest, to an extent and with similar statistical significance to what is seen with 2Mbps. As dramatic example in the other direction, the outcome for services in the 150-155Mbps range is a coefficient of +0.378 with $t = 11.2$. These services are on average priced *higher* than the model, to a much larger extent than the 2Mbps services are priced lower and with a deviation that is, if anything, even more statistically significant. Undoubtedly, the buyers of 100Mbps services and the sellers of 150-155Mbps services would also prefer their products to be priced separately from the model.
- 6.5 A second problem with Models 4 and 5 is they introduce an arbitrary break point between two groups of services and a discontinuity in pricing between the two groups. EI clearly understand the second part of this concern, with their illustration of a price jump from \$341 to \$431 for a miniscule change in specifications (p.54). The definition of precisely which services are singled out for special treatment is also arbitrary: Why

[C-i-C begins] = information not to be released without a confidentiality undertaking
[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

2.5Mbps as the capacity break point? Why 5km as the distance break point? Is it just metro routes or all routes meeting the criteria?

- 6.6 The pricing formulas derived from these models will also be more complex than with the equivalent models in the EI Final report because of the need to specify the separate pricing within the different groups of services.
- 6.7 Model 6 removes the problem of the discontinuity by providing a ramp up from the break point instead of a sudden jump, but at the expense of further complexity. Not only does it appear the precise meaning of “2Mbps” needs to be different to accommodate the technicalities of Model 6, the resulting pricing formula will be more complex again than those from Models 4 or 5 due to the additional nonlinearity.
- 6.8 The observation by Optus – that the benchmarking models fitted to data from competitive routes will predict higher regulated prices on some declared routes (particularly low-capacity/short-distance ones) than the historical prices available to Optus on similar routes – does not invalidate the benchmarking process. It is no surprise that a large user of these services such as Optus has been able to negotiate a commercial deal for bundles of services across exempt and declared routes at average prices well below the competitive prices available to one-off customers. Indeed, if the Optus acquisitions in this bulk purchase were to be preponderantly on declared routes, we would see a situation where the benchmark data did not reflect the favourable prices available when the bulk discounts are taken, resulting in exactly the situation Optus observes. The ACCC has insisted that commercial discounts, such as might be available for large purchases, will continue to be available in the regulated market for DTCS services. It is not the intention in benchmarking that the very best discounts that have been available to the largest customers should become the regulated prices specified by the FAD. For that reason the whole motivation for Chapter 4 (and much else in the EI Further report) is derived from a misunderstanding.

7. Stochastic frontier analysis (EI Further report, Chapter 5)

- 7.1 The attraction to some stakeholders in using stochastic frontier analysis (SFA) instead of regression modelling to establish the benchmark relationship between the prices on exempt routes and the driver variables appears to come from the low prices predicted by SFA. Low predicted prices is not a surprising outcome, because SFA seeks to fit an envelope to a *lower bound* on the variable prices at each combination of attributes, rather than to the average of such variation as in regression models. The bound is stochastic to allow for a degree of symmetric noise in the data, while the one-sided component of the error term allows all variations in prices upward from the lower bound to be interpreted

[C-i-C begins] = information not to be released without a confidentiality undertaking

[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

as ‘inefficiencies’ – and re-interpreted in this application as non-competitive additions to lowest-cost pricing.

7.2 The unsuitability of SFA for benchmarking in the DTCS was effectively dealt with in EI Final report and my response to that report, so it is surprising that the method should resurface here. The main arguments against this method for benchmarking were identified earlier as follows:

- (i) “the cross-sectional SFA model discussed and apparently tested by the expert who made this recommendation did not appear to be consistent with the economic arguments it put forward, which tended to suggest that due to widespread bundling practices, some of the providers may retain some degree of market power.” (EI Final report, p.88);
- (ii) “A problem with the SFA approach in this context is that it would forecast lower prices based on an efficiency interpretation of the unexplained variation in the data, but given the scope of this variation, a premium would then need to be added to ensure prices were sufficient to finance investment and allow for estimation uncertainty. But it is not clear what the premium should be or how to calculate it.” (EI Final report, p.44); and
- (iii) “...stochastic frontier models proposed by other experts do not answer the fundamental question of benchmarking against average competitive pricing.” (Breusch on Final, paragraph 3.5)

7.3 It is ironic that bundling should be used as a motive for the use of SFA in benchmarking, when price averaging in bundling works to undermine the interpretation that is given to the SFA results. If services with widely different characteristics and hence very different cost bases are sold at the same common price, the services that are more costly to produce will be accorded considerable cross-subsidies. ([C-i-C begins] [Confidential] [C-i-C ends]) However, SFA will use that very low observed price of the highly subsidised service as an important data point in establishing the lower bound envelope of ‘efficient’ prices. Inevitably, SFA will understate the minimum cost of production in such cases.

7.4 I have not attempted to replicate the results involving SFA in the EI Further report, Chapter 5. EI’s finding that the result is lower predicted prices in all market segments is not at all surprising. The unsuitability of the method for benchmarking the DTCS is unchanged by any of these new estimation or prediction results.

[C-i-C begins] = information not to be released without a confidentiality undertaking
[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

8. Issues carrying over from the EI Final report/Draft FAD

Productivity movements over time using ‘contract start date’

- 8.1 Model 1 of the EI Final report includes the variable ‘contract start date’, and section 6.2.3 of that report proposes using the corresponding regression coefficient to infer a rate of price decline over time to represent productivity gains in supplying the DTCS. Model 1 is not a preferred model in the EI Final report, and the implications of that model are not developed into a pricing equation, in part because of advice from providers that the records of contract start dates do not properly account for contract renewals and hence are not informative of point-in-time pricing. As EI correctly state, there is a “lack of information to provide a robust empirical basis for specifying a productivity adjustment factor” (EI Final report, p.61). In Breusch on Final, I speculate that another reason this approach was not proceeded with is the small rate of price decline implied by Model 1, which is less than two per cent a year.
- 8.2 Despite the conclusions reached in the EI Final report, the variable ‘contract start date’ appears in all models with suffix ‘a’ in the EI Further report. In Model 1a the implication is a rate of price decline of slightly more than 2 per cent a year, while in Model A1a it is again slightly less than 2 per cent a year. Aside from the magnitude of the adjustment, the earlier concerns with the quality of this variable, and the “lack of information to provide a robust empirical basis” for the adjustment factor, remain the same as before. For that reason, the models with suffix ‘a’ should not be used for pricing the DTCS.

Route and ESA throughput variables

- 8.3 Model 2 of the EI Final report includes the variables ‘route throughput’ and ‘ESA throughput’, intended to capture economies of scale that may be available at busier ESAs or routes. In stating a preference for Model 2 over Model 3, which excludes these throughput variables, the Draft FAD states:

The ACCC considers [the pricing formula derived from Model 2] the most appropriate model for setting regulated prices as it recognises that regulated routes typically have lower throughput than competitive routes. The model accounts for the different economies of scale in regulated routes through the route throughput and ESA throughput variables. (Draft FAD, p.30)

- 8.4 Unfortunately, the coefficient on ESA throughput – as found in both Models 1 and 2 of the EI Final report as well as in all models with suffix ‘a’ and ‘b’ of the EI Further report – turns out to be positive, where a negative coefficient is predicted. Thus the empirical evidence is inconsistent with the hypothesis of scale economies as advanced in the Draft FAD. The EI Final report attempts to cover this unfortunate outcome with an alternative hypothesis of capacity constraints, although that is only an ex-post rationalisation of an

[C-i-C begins] = information not to be released without a confidentiality undertaking

[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

unexpected outcome and lacks any logical basis. In Breusch on Final, I proposed that this logical hurdle could be avoided by excluding the throughput variables, which in any case make little contribution to the statistical fit while at the same time adding unnecessary complexity to the pricing formula. In the context of the EI Further report, the implication is that models with suffix 'a' or 'b' should not be used for pricing the DTCS.

2km distance assumption for tail-ends

8.5 The price predictions in the EI Further report continue to use the earlier assumption of a nominal distance of 2 kilometres for each tail-end service. The 2km assumption appears to be supported by some analysis in section 6.3 of the EI Final report. However, in Breusch on Final, I provide several reasons why EI's analysis must produce an *under*-estimate of the average distance: the assumption of a circular ESA has the smallest average radial distance of any possible ESA shape; every ESA cannot be circular because that would not give coverage (and no real ESA is circular); assuming the exchange is exactly at the centre of the circular ESA further minimizes the radial distance; the average radial distance to points that are evenly distributed within a circle is in fact *two-thirds* of the radius to the edge of the circle, not one-half as stated by EI. In an alternative calculation, where assumption of a circular ESA with the exchange at its centre is made for ESAs *on average* rather than so unrealistically for each ESA, the average radial distance is 3.51km.

Connection charges

8.6 The Draft FAD sets connection charges in the DTCS independently of the regression models developed by EI, using a separate benchmarking process. It is not entirely clear from the Draft FAD or the EI Final report how this separate benchmarking is to be done, although it seems the method uses averages of connection charges observed in the data file, where those averages are calculated separately within groups of services classified by capacity and interface (presumably only for services on exempt routes, although that is not clear, either). In Breusch on Final, I noted several statistical challenges to this process due to missing data, extreme outliers and small samples. With unclear descriptions and these unanswered problems of statistical method, the process for setting connection charges by benchmarking is far from transparent.

Ethernet versus SDH

8.7 The Draft FAD canvasses the possibility of pricing all services as if they are Ethernet, even when the actual technology is SDH:

The ACCC would also like stakeholders' views on whether interface type should be allowed to vary or whether to fix interface type at Ethernet (zero). There may be some merit in setting the regulated price on the basis of Ethernet as the interface type as

[C-i-C begins] = information not to be released without a confidentiality undertaking

[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

Ethernet is the newer technology and is increasingly used in preference to SDH.

Ethernet may also be more efficient and cost effective than SDH. (Draft FAD, p.31)

The estimated coefficient of the SDH indicator is positive, representing higher average prices for SDH services relative to Ethernet by 20-30 per cent, depending on the model. The ACCC's proposal will depress the regulated prices of SDH services relative to the benchmark by the percentage of the estimated coefficient.

- 8.8 The difficulty with the proposal is it assumes that the higher selling price for SDH services is due solely to higher production costs of that technology. But these are all services on exempt routes, where the markets are deemed to be competitive, so any cost variations in providing a standard product would be borne by the seller not the buyer. The higher production costs hypothesis would explain why SDH is being replaced with Ethernet, but it does not explain why SDH services continue to be sold at premium prices on exempt routes. Perhaps the SDH variable is representing other factors of service quality that are not explicitly included in the model. Without a clear economic rationale for the estimated coefficient, it is wrong to rely on the interpretation of cost effectiveness to manipulate the pricing outcomes in this way.

9. Overview of benchmarking the 2015 DTCS FAD.

- 9.1 In the early stages of the econometric work for the 2015 DTCS FAD, I was critical of the EI Workshop paper and Draft report because the approach was unduly reliant on criteria that are more appropriate to statistical inference in an academic context, with too much speculative theorising, and with insufficient attention to the requirements of a benchmark pricing equation. All of these inappropriate directions improved with the EI Final report, to the extent that a workable, fair and transparent outcome seemed close.
- 9.2 Now the EI Further report has opened up issues that seemed to be finalised by the "Final" report. So Model 1 with its doubtful variable 'contract start date' is back on the table for consideration in the form of the models with suffix 'a'. The three old issues of outliers, 2Mbps services and SFA have been revived after being largely settled in the Final report.
- 9.3 The EI Further report not only rekindles matters that seemed to be settled, but in doing so it uses criteria in the model selection process that go beyond benchmarking. The most obvious of these new found requirements is that the model should predict prices on low-capacity metro routes (the so-called 2Mbps services) lower than the prices Optus has been able to obtain in the past.
- 9.4 After more than 370 pages in four stages of consultant reports from EI (not counting the re-issues with corrections), several hundred files of EI's computer code and spreadsheets of intermediate calculations, not to mention the data files distributed to experts by the

[C-i-C begins] = information not to be released without a confidentiality undertaking
[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

ACCC and the various reports prepared by those experts, it is not clear that the process of securing a model for benchmarking the DTCS has yet converged. Indeed, the developments since September 2015 have been outwards.

- 9.5 It is salutary to compare the DTCS FAD 2015 with the process for the FAD 2012. The modelling work undertaken by EI in 2015 has been many times more detailed than that by DAA in 2012. It is an open question how much the outcome has improved, if at all. Even within the four stages of EI's own reports (Workshop, Draft, Final, Further), there has been growth, with each report being longer than the previous one.
- 9.6 This burgeoning complexity and detail must be curbed if benchmarking is to remain a viable method of price regulation into the future. Of course, some stakeholders may prefer to have prices set another way, irrespective of the resource cost of doing so, but that would not be in the public interest.

10. Conclusions

- 10.1 The EI Further report confirms that the models of the EI Final report provide a robust basis for benchmarking the DTCS. There is clear evidence that the additional JV data are different from the original data and should not be used when benchmarking the DTCS against competitive prices.
- 10.2 The original data contained some errors relating to GST, so the models of the EI Final report need to be re-estimated, giving Models A1a, A1b and A1c as the replacements for Models 1, 2 and 3, respectively. In the EI Final report, both Models 2 and 3 are described equally as "preferred" and both are developed through to a pricing formula. Interestingly, while Model 2 is adopted for the Draft FAD, there is a slip in the EI Further report that suggests the consultants in fact preferred Model 3 above Model 2. (EI Further report, p.5)
- 10.3 In Breusch on Final, I noted the strong case for Model 3 over Model 2, based on economic logic, statistical fit and the simplicity and transparency of the resulting pricing formula. All of these same criteria remain in place to indicate a preference for Model A1c over Model A1b (and A1a). In fact, that preference is stronger than before, because on statistical fit measured by the BIC, Model A1c is now clearly better than Model A1b. (EI Further report, Table A.2) The pricing formula for Model 3 in the EI Final report is easily updated with the new coefficient values, giving a principled and robust outcome to the 2015 DTCS FAD.

[C-i-C begins] = information not to be released without a confidentiality undertaking
[C-i-C begins] = information not to be released even with a confidentiality undertaking

Public version

11. References

Australian Competition and Consumer Commission. *Final Access Determination No.1 of 2012 (DTCS)*, July 2012. (“**2012 FAD**”)

Australian Competition and Consumer Commission. *TFA - 2015 DTCS FAD - May 2015 - revised data set for EI and experts - CIC commercial in confidence copy.xlsx*. Computer file downloaded 12 May 2015. (“**original data**”)

Australian Competition and Consumer Commission. *Public Inquiry to make a Final Access Determination for the Domestic Transmission Capacity Service: Draft Decision*. 4 September 2015. (“**Draft FAD**”)

Australian Competition and Consumer Commission. *2015 DTCS FAD - revised data set including Optus VHA eJV joint venture pricing - commercial in confidence CIC v6 - GST de-identified for experts.xlsx*. Computer file downloaded 23 December 2015. (“**ACCC data file**”)

Breusch, Trevor. *Report on: Economic Insights Working Paper prepared for the ACCC in relation to the 2015 DTCS FAD*, May 2015. (“**Breusch on Workshop**”)

Breusch, Trevor. *Report on: Economic Insights Draft Report prepared for the ACCC in relation to the 2015 DTCS FAD*, 7 July 2015. (“**Breusch on Draft**”)

Breusch, Trevor. *Report on: Economic Insights Final Report and 2015 DTCS Draft FAD*, 2 October 2015. (“**Breusch on Final**”)

Economic Insights. *Domestic Transmission Capacity Services Benchmarking Model, Workshop Paper prepared for Australian Competition and Consumer Commission*, authors John Fallon, Michael Cunningham, and Tim Coelli, 21 April 2015. (“**EI Workshop paper**”)

Economic Insights. *Domestic Transmission Capacity Services Benchmarking Model: Draft Report prepared for Australian Competition and Consumer Commission*, authors John Fallon, Michael Cunningham, and Tim Coelli, 11 June 2015. (“**EI Draft report**”)

Economic Insights. *Domestic Transmission Capacity Services Benchmarking Model: Final Report prepared for Australian Competition and Consumer Commission*, authors John Fallon, Michael Cunningham, and Tim Coelli, 1 September 2015. (“**EI Final report**”)

Economic Insights. *Domestic Transmission Capacity Services Benchmarking Model: Testing Further Specifications: Report prepared for Australian Competition and Consumer Commission*, authors John Fallon, Michael Cunningham, and Tim Coelli, 16 December 2015. (“**EI Further report**”)
