

QUANTITATIVE ECONOMICS

Memorandum

1 Introduction

This memo examines options for further development of the opex cost function used in benchmarking distribution network service providers (DNSP). The aims of this exercise are to develop options for dealing with the issue of monotonicity violations in translog (TLG) opex cost function models, and to identify areas for further investigation with regard to the specification or use of opex cost function results. It includes a brief discussion of averaging efficiency scores.

1.1 Motivation

The key issue motivating this investigation the presence of an undue number of monotonicity violations in some the translog opex cost function models. In the TLG specification, the elasticity of real variable cost with respect to each output varies with different levels of all the outputs (unlike the Cobb-Douglas model in which they are constant). The monotonicity condition requires that these elasticities of cost with respect to each output should be positive for all the observations. If there is a large number of violations, this warrants further investigation. Kumbhakar, Wang and Horncastle (2015, 107) suggest that a "number of factors may lead to such violations. Imposing more structure in the estimation process … could make the results align more with the theory". This suggestion motivates the approach taken in this memo.

It remains a question as to whether the monotonicity violations can be entirely interpreted as such. A short-run variable cost function typically includes a measure of the quasi-fixed capital stock as an explanatory variable. Economic Insights (2014) chose not to include a capital input variable, partly due to the data limitation of obtaining a consistent measure across jurisdictions, and partly due to the statistical limitation of a high degree of correlation between the capital input measure and the output variables. This correlation means that the omission of capital input, while not biasing the measures of opex efficiency, may have implications for inferences concerning the output elasticities, including the interpretation of instances of negative output elasticities as necessarily reflecting only a monotonicity violation. This question may be further investigated in future opex cost function development work.

1.2 Considerations for evaluating models

The models presented here are compared against the base model, which is that presented in the draft DNSP benchmarking report and reproduced for convenience in Appendix A. The main considerations in evaluating alternative models are:

- the extent to which it reduces monotonicity errors for Australian DNSPs
- joint significance tests of groups of related explanatory variables added to a model
- goodness-of-fit, and
- the meaningful economic interpretation of parameters.

For some of the estimation methods used here, conventional goodness-of-fit statistics are not available. To provide a common basis of comparison, we measure goodness-of-fit using the pseudo adjusted $R²$ statistic, defined as:

$$
adj. R^2 = 1 - (1 - r^2) \left(\frac{N - 1}{N - k}\right) \tag{1}
$$

where *r* is the correlation coefficient between the actual and predicted values of the dependent variable; *N* is the number of observations; and *k* is the number of parameters in the model (including the intercept). This measure of fit penalises the addition of unnecessary explanatory variables. For example, the Cobb-Douglas (CD) model is more parsimonious than the TLG model and the adjusted \mathbb{R}^2 will tend to favour the former unless the additional explanatory variables in the TLG model have a strong influence.

1.3 Some statistics for the base model

1.3.1 Monotonicity violations

Referring only to the base TLG models in Appendix A, a summary is presented in Table 1.1. For the sample period 2006 to 2021, both the LSETLG and SFATLG models produce monotonicity violations in just over 12 per cent of all observation and no monotonicity violations for observations on Australian DNSPs. This is a suitable outcome.

For the sample period 2012 to 2021, both the LSETLG and SFATLG models produce more monotonicity violations than for the longer sample period. Considering the proportion of observations for which there are monotonicity violations for *any* of the outputs, there are 28.8 per cent for the LSETLG model and 43.8 per cent for the SFATLG model over all observations and any output. For the Australian DNSPs' observations, the monotonicity

violations account for 42.3 per cent for the LSETLG model and 65.4 per cent for the SFATLG model. For the LSETLG model, five Australian DNSPs have monotonicity violations in more than 50 per cent of all observations, and for the SFATLG model nine have monotonicity violations in more than 50 per cent of all observations. This is not a satisfactory outcome.

1.3.2 Goodness-of-fit

The goodness-of-fit measures are all quite high for models estimated using the 2006 to 2021 sample, and the corresponding fit measures for the 2012 to 2021 period are on average at a similar level. Focussing on the base models for the period 2006 to 2021 only (see Appendix A), the LSECD model has $adj.R^2 = 0.977$, which is slightly lower than the LSETLG model with $adi. R^2 = 0.980$. The SFACD model has $adi. R^2 = 0.969$, which is slightly higher than the LSETLG model with *adj.R2* = 0.961. The average of the CD models' (LSECD and SFACD) adjusted R^2 is 0.973. The average of the TLG models' adjusted R^2 is 0.971. This suggests that, on average, the TLG models do not improve on the fit of the CD models.

1.3.3 Joint significance tests

The Base TLG specification has six additional parameters compared to the CD model, and we can test the statistical significance of their inclusion.

- *Using the 2006 to 2021 sample*: The null hypothesis of the Wald test of the higher-order coefficients of the LSETLG collectively being zero is rejected (Prob $> \chi^2 = 0.0000$). The null hypothesis of the Wald test of the higher-order coefficients of the SFATLG collectively being zero fails to be rejected (Prob $> \chi^2 = 0.4440$).
- *Using the 2012 to 2021 sample*: The null hypothesis of the Wald test of the higher-order coefficients of the LSETLG collectively being zero is rejected indicating that the higher-order coefficients are non-zero (Prob $> \chi^2 = 0.000$). The null hypothesis of Wald

Opex cost function development

test on the higher-order coefficients of the SFATLG model is rejected (Prob $> \chi^2 =$ 0.0001).

The first of these results indicates that when the full sample is used the retained higher-order terms are only jointly significantly different from zero in one of the two versions of the model.

2 Restricting the higher-order and interaction effects in the TLG models

2.1 Approach

Economic Insights (2021, 140–41), examined some variations to the TLG specification, the most promising of which involved omitting some of the higher-order or interaction effects (called 'hybrid' models). It was found that a specification in which the squared terms applying to the two highly correlated outputs (customer numbers and RMD) are excluded, but the squared term on circuit length and the three interaction terms are retained reduced the frequency of the monotonicity violations for Australian DNSPs in various different sample periods tested. The remaining 2nd-order terms were, depending on the sample period, either jointly significant in both the LSE and SFA models, or jointly significant in at least one of those models. The fact that these terms were not jointly significant in all cases was a limitation of that specification. This particular hybrid specification is tested here with an additional year of data to see if it performs adequately in the larger sample. For our purposes this model is called 'TLG-H1'.

Another 'hybrid' model, which was not tested previously, is tested here, based on the hypothesis that, whilst there may be economies of scale when more energy is delivered through a fixed network, there are not necessarily economies of scale in the spatial extension of the energy network. Hence, this variation removes the higher-order terms relating to circuit length. This model will be referred to as 'TLG-H2'.

A third hybrid model (referred to as 'TLG-H3') does not include any squared log outputs and the only included higher-order terms are the interaction between customer numbers and RMD and the interaction between circuit length and RMD.

Table 2.1 summarises the variable specifications of the three hybrid models compared to the base models.

Table 2.1 Summary of model specifications tested

2.2 Results

The results of estimating the two models discussed above are presented in Appendix B.

2.2.1 TLG-H1 model

Table 2.2 shows a summary of the monotonicity errors for the TLG-H1 model. Both the LSE and SFA versions of this model produce fewer monotonicity violations over all observations, when the 2006 to 2021 sample is used. Between 2 and 3 per cent of observations compared to just over 12 per cent for the Base TLG model. They also both produce zero monotonicity violations for the Australian DNSPs, the same as the Base TLG models.

When the 2012 to 2021 sample is used, these models also improve on the monotonicity results of the Base TLG model. For the LSE H1-TLG model only one, and for the SFA TLG-H1 model only two, Australian DNSP have monotonicity violations for more than 50 per cent of their observations.

Table 2.2 TLG-H1: Frequency of monotonicity violations: Summary

As pointed out in section 1.3, the goodness-of-fit of the Base CD and TLG models are broadly similar. Since the fit of TLG-H1 models (which are an intermediate specification) is again similar. The TLG-H1 specification has four additional parameters compared to the CD model, and we can test the statistical significance of their inclusion.

- *Using the 2006 to 2021 sample*: The null hypothesis of the Wald test of the higher-order coefficients of the LSETLG collectively being zero is rejected (Prob $> \chi^2 = 0.0001$). The null hypothesis of the Wald test of the higher-order coefficients of the SFATLG collectively being zero fails to be rejected (Prob $> \chi^2 = 0.2905$).
- *Using the 2012 to 2021 sample*: The null hypothesis of the Wald test of the higher-order coefficients of the LSETLG collectively being zero is rejected indicating that the higher-order coefficients are non-zero (Prob $> \chi^2 = 0.000$). The null hypothesis of Wald test on the higher-order coefficients of the SFATLG model is rejected (Prob $> \chi^2 =$ 0.0140).

The first of these results indicates that the full sample is used the retained higher order terms are only jointly significantly different from zero in one of the two versions of the model.

2.2.2 TLG-H2 model

Table 2.3 shows a summary of the monotonicity errors for the TLG-H2 model. Once again, there are fewer monotonicity violations over all observations, when the 2006 to 2021 sample is used. There are no monotonicity violations for the Australian DNSPs using this sample period, which is the same as for the Base TLG models. The number of monotonicity violations over all observations is less that the Base models, and is zero for the SFA version of TLG-H2.

	LSETLG-H2		<i>SFATLG-H2</i>	
	All obs	Aust, obs	All obs	Aust. obs
2006 to 2021				
Customer numbers	0.0%	0.0%	0.0%	0.0%
Circuit length	0.0%	0.0%	0.0%	0.0%
RMD	8.9%	0.0%	0.0%	0.0%
Total	8.9%	0.0%	0.0%	0.0%
2012 to 2021				
Customer numbers	0.5%	0.0%	7.1%	19.2%
Circuit length	0.0%	0.0%	0.0%	0.0%
RMD	11.3%	0.0%	8.6%	0.0%
Total	11.7%	0.0%	15.6%	19.2%

Table 2.3 TLG-H2: Frequency of monotonicity violations: Summary

When the 2012 to 2021 sample is used, the TLG-H2 models improve significantly on the monotonicity results of the Base TLG model. For the LSE version of this model, there are no monotonicity violations for the Australian DNSPs compared to 42.7 per cent for the Base LSETLG model. For the SFA version, there are monotonicity violations in 19.2 per cent of the observations for the Australian DNSPs compared to 65.3 per cent for the Base SFATLG model. Only two Australian DNSP have monotonicity violations for more than 50 per cent of their observations using the SFA version of the TLG-H2 specification.

The TLG-H2 specification has three additional parameters compared to the CD model, and we can test the statistical significance of their inclusion.

- *Using the 2006 to 2021 sample*: The null hypothesis of the Wald test of the higher-order coefficients of the LSETLG collectively being zero is rejected (Prob $> \chi^2 = 0.0000$). The null hypothesis of the Wald test of the higher-order coefficients of the SFATLG collectively being zero fails to be rejected (Prob $> \chi^2 = 0.4414$).
- *Using the 2012 to 2021 sample*: The null hypothesis of the Wald test of the higher-order coefficients of the LSETLG collectively being zero is rejected indicating that the higher-order coefficients are non-zero (Prob $>\chi^2$ = 0.000). The null hypothesis of Wald test on the higher-order coefficients of the SFATLG model is rejected (Prob $> \chi^2 =$ 0.0019).

Like the TLG-H1 specification, the first of these indicates when the full sample is used, that the retained higher order terms are only jointly significantly different from zero in one of the two versions of the model.

2.2.3 TLG-H3 model

Table 2.4 shows a summary of the monotonicity errors for the TLG-H3 model. Once again, there are fewer monotonicity violations over all observations, when the 2006 to 2021 sample is used. There are no monotonicity violations for the Australian DNSPs using this sample period, which is the same as for the Base TLG models. The number of monotonicity violations over all observations is less that the Base models, and is zero for the LSE version of TLG-H3.

When the 2012 to 2021 sample is used, the TLG-H3 models improve considerably on the monotonicity results of the Base TLG model. For the LSE version of this model, there are no monotonicity violations for the Australian DNSPs in either the LSE or SFA versions of this model.

The TLG-H3 specification has two additional parameters compared to the CD model, and we can test the statistical significance of their inclusion.

• *Using the 2006 to 2021 sample*: The null hypothesis of the Wald test of the higher-order coefficients of the LSETLG collectively being zero is rejected (Prob $> \chi^2 = 0.0003$). The null hypothesis of the Wald test of the higher-order coefficients of the SFATLG collectively being zero fails to be rejected (Prob $> \chi^2 = 0.0971$).

• *Using the 2012 to 2021 sample*: The null hypothesis of the Wald test of the higher-order coefficients of the LSETLG collectively being zero is rejected indicating that the higher-order coefficients are non-zero (Prob $>\chi^2$ = 0.001). The null hypothesis of Wald test on the higher-order coefficients of the SFATLG model is rejected (Prob $> \chi^2 =$ 0.6786).

Quantonomics

QUANTITATIVE ECONOMICS

Like the other two hybrid specifications, the retained higher order terms are only jointly significantly different from zero in one of the two versions of the model, in this case in both the full sample and the sample from 2012 to 2021.

Table 2.4 TLG-H3: Frequency of monotonicity violations: Summary

2.3 Summary conclusion

Although the restricted models considered in this section reduce the number of monotonicity violations, in some cases the additional higher-order terms of the model are not jointly statistically significant. This must be regarded as a limitation of these hybrid models.

For each specification, there are four models, the LSE and SFA variants, and two sample periods, 2006 to 2021 and 2012 to 2021. Averaging over the four models, we find the average number of monotonicity violations for Australian DNSPs is: (i) 5.8 per cent for the H1 models, (ii) 4.8 per cent for the H2 models, and (iii) zero for the H3 models. The average adjusted \mathbb{R}^2 is: (i) 0.974 for H1, (ii) 0.975 for H2, and (iii) 0.973 for H3.

The H2 specification has an economic rationale. A number of studies, including Roberts (1986) and Llorca et al (2016), have found that while there are usually economies of scale in expanding the volume of energy conveyed over a fixed network, they find no economies of scale in extension of the area supplied (total line length). The H2 specification is loosely based on this idea. The H2 specification provides a (very) slightly better fit than the other two models and keeps monotonicity violations to a reasonably low level.

That said, these alternative models need to be evaluated in light of their proposed application. In circumstances where the TLG models have too many monotonicity violations, a hybrid model could be used in its place when calculating the average of four econometric models. The H3 model produces no monotonicity violations for Australian DNSPs, and would therefore be usable in more circumstances.

3 Alternative estimation methods

We have also explored an alternative method of stochastic frontier estimation which uses robust regression methods. Robust regression is a method which reduces the weight given to extreme outliers which are otherwise overly influential on parameter estimation. The robust SFA estimation method we tested has been used by Wheat, Stead and Greene (2019), and implemented in a Stata community contributed program *rfrontier* produced by Alexander Stead.¹ The model we tested differs in another important respect from the SFA methods used in the Base and other models presented here. The inefficiency terms are assumed to have a half-normal distribution rather than a truncated normal distribution. The robust SFA estimator we tested assumes that the stochastic residual has a Student's *t* distribution, and the 'true' underlying stochastic disturbance is Normally distributed.

Robust regression methods could presumably also be used as an alternative to the least-squares estimator (LSE), with the most comparable methods being an M-estimator or MM-estimator which use maximum likelihood methods. This approach has not been tested. We would need to undertake further investigation into the statistical properties of the efficiency scores obtained via fixed effects within the context of a robust regression method.

We do not report the detailed SFATLG robust regression results in this memo, but highlight the following preliminary findings:

- It produced no monotonicity violations of Australian DNSPs in either the 2006 to 2021 sample or in the 2012 to 2021 data sample. The only monotonicity violations were for NZ DNSPs and related to the RMD output.
- The higher-order terms were jointly statistically significant using either the 2006 to 2021 sample, or the 2012 to 2021 sample.

These results suggest that the robust SFA method may be a promising area for investigation. They may suggest that extreme outliers have some influence on the incidence of monotonicity violations. That said, we note that in regard to robust regression in general, there are several different methods, and often require judgement in the use of 'tuning parameters', and the specific preferred method may be difficult to justify in a contested context. In regard to the

¹ https://github.com/AlexStead/rfrontier.

SFA robust regression method in particular, we note many issues that would limit its application to the opex cost function, at least at present:

- The *rfrontier* program does not have the option of using the truncated normal distribution for inefficiency terms, which is the preferred assumption used to date in SFA analysis. The half-normal distribution assumption is more restrictive.
- It is a quite new method and hence not as thoroughly tried and proven a technique as would be necessary for a method used in a regulatory application.

4 Conclusions and comment on averaging efficiency scores

The analysis of hybrid model options has been examined and some feasible options have been identified. The TLG-H1 specification substantially reduces the number of monotonicity violations compared to the base model. The TLG-H2, which removes the higher-order terms relating to circuit length, reduces the monotonicity violations on balance more than the TLG-H1 specification, and it has no monotonicity violations for Australian DNSPs in the LSETLG-H2 model using the sample 2012 to 2021. The TLG-H3 specification, which has no squared log outputs and includes interactions between customer numbers and RMD and between circuit length and RMD, has no monotonicity violations for Australian DNSPs in both the LSE and SFA versions using the sample 2012 to 2021.

In circumstances where the Base TLG models have too many monotonicity violations, one of these hybrid models could be used in its place when calculating the average of four econometric models. When making such a substitution, a model with the corresponding estimation technique should be used. For example, if the Base LSETLG model has too many monotonicity violations, the efficiency score obtained from the LSE version of the hybrid model should be used in its place (eg, LSETLG-H3). An efficiency score from an SFA hybrid model should not be substituted for an efficiency score from the LSETLG model, and vice versa.

An approach based on 'robust regression' has been tested. Whilst this is a promising line of investigation in future, it would be premature to use this method at the present time.

Appendix A: Base Model

The models presented here are the same as those presented in Appendix C of the Draft Report.

A1 Sample 2006 to 2021

Table A.1 Base CD: DNSP Opex cost function (2006–2021)

QUANTITATIVE ECONOMICS

Table A.2 Base TLG: DNSP Opex cost function (2006–2021)

Table A.3 Base TLG: Average DNSP output elasticities by country (2006–2021)

Table A.4 Base TLG: Frequency of monotonicity violations by country (2006–2021)

Table A.5 Base TLG: Frequency of monotonicity violations by Aust. DNSP (2006-2021)

A2 Sample 2012 to 2021

Table A.6 Base CD: DNSP Opex cost function (2012–2021)

Table A.7 Base TLG: DNSP Opex cost function (2012–2021)

Table A.8 Base TLG: Average DNSP output elasticities by country (2012–2021)

Table A.9 Base TLG: Frequency of monotonicity violations by country (2012–2021)

Table A.10 Base TLG: Frequency of monotonicity violations by Aust. DNSP (2012-2021)

Appendix B: Restricted or Modified Translog Models

B1 TLG-H1

B1.1 Sample 2006 to 2021

Table B.1 TLG-H1: DNSP Opex cost function (2006–2021)

Table B.2 TLG-H1: Average DNSP output elasticities by country (2006–2021)

Table B.3 TLG-H1: Frequency of monotonicity violations by country (2006–2021)

Table B.4 TLG-H1: Frequency of monotonicity violations by Aust. DNSP (2006-2021)

B1.2 Sample 2012 to 2021

Table B.5 TLG-H1: DNSP Opex cost function (2012–2021)

Table B.6 TLG-H1: Average DNSP output elasticities by country (2012–2021)

Table B.7 TLG-H1: Frequency of monotonicity violations by country (2012–2021)

Table B.8 TLG-H1: Frequency of monotonicity violations by Aust. DNSP (2012-2021)

B2 TLG-H2

B2.1 Sample 2006 to 2021

Table B.9 TLG-H2: DNSP Opex cost function (2006–2021)

Table B.10 TLG-H2: Average DNSP output elasticities by country (2006–2021)

Table B.11 TLG-H2: Frequency of monotonicity violations by country (2006–2021)

Table B.12 TLG-H2: Frequency of monotonicity violations by Aust. DNSP (2006-2021)

B2.2 Sample 2012 to 2021

Table B.13 TLG-H2: DNSP Opex cost function (2012–2021)

Table B.14 TLG-H2: Average DNSP output elasticities by country (2012–2021)

Table B.15 TLG-H2: Frequency of monotonicity violations by country (2012–2021)

Table B.16 TLG-H2: Frequency of monotonicity violations by Aust. DNSP (2012-2021)

B3 TLG-H3

B3.1 Sample 2006 to 2021

Table B.17 TLG-H3: DNSP Opex cost function (2006–2021)

Table B.18 TLG-H3: Average DNSP output elasticities by country (2006–2021)

Table B.19 TLG-H3: Frequency of monotonicity violations by country (2006–2021)

Table B.20 TLG-H3: Frequency of monotonicity violations by Aust. DNSP (2006-2021)

B3.2 Sample 2012 to 2021

Table B.22 TLG-H3: Average DNSP output elasticities by country (2012–2021)

Table B.23 TLG-H3: Frequency of monotonicity violations by country (2012–2021)

References

- Economic Insights. 2014. 'Economic Benchmarking Assessment of Operating Expenditure for NSW and ACT Electricity DNSPs'. Report Prepared for Australian Energy Regulator by Denis Lawrence, Tim Coelli and John Kain.
	- ———. 2021. 'Economic Benchmarking Results for the Australian Energy Regulator's 2021 DNSP Annual Benchmarking Report'. Draft Report Prepared by Michael Cunningham, Denis Lawrence and Tim Coelli for Australian Energy Regulator.
- Kumbhakar, Subal C., Hung-Jen Wang, and Alan P. Horncastle. 2015. *A Practitioner's Guide to Stochastic Frontier Analysis Using Stata*. Cambridge University Press.
- Llorca, Manuel, Luis Orea, and Michael G. Pollitt. 2016. 'Efficiency and Environmental Factors in the US Electricity Transmission Industry'. *Energy Economics* 55 (March): 234–46. https://doi.org/10.1016/j.eneco.2016.02.004.
- Roberts, Mark J. 1986. 'Economies of Density and Size in the Production and Delivery of Electric Power'. *Land Economics* 62: 378–87.
- Wheat, Phill, Alexander D. Stead, and William H. Greene. 2019. 'Robust Stochastic Frontier Analysis: A Student's t-Half Normal Model with Application to Highway Maintenance Costs in England'. *Journal of Productivity Analysis* 51 (1): 21–38.