

Note on opex modelling in the DPP4 Issues paper

30 November 2023

The [DPP4 Issues paper](#) published 2 November 2023 (the paper) includes opex model results. The results in tables D6 and D7 differ to the results generated by running the [code](#) we published.

The data and regression fitting methods are the same. The results in the paper reflect some data treatments varied from the published code. This stems from our exploration of aspects of the modelling which we note in the paper we need to determine.

This note presents these differences. The accompanying R script generates tables (shown below) like D6 and D7 using the data treatment in the published code, while allowing these to be varied. We are providing these to support submissions on this section of the paper. We have no changes to the conclusions or questions in the paper.

We still need to determine a range of modelling aspects

In the paper we note aspects of our econometric modelling we need to determine, including time-period, exclusion of outliers, and cost escalators used to deflate historical costs (para. D15). Our exploration of these data treatments lies behind the differences between the values shown in tables D6 and D7 and the published code output.

Tables D6 and D7 were generated in extension to the published code. We like that they show both model fit metrics and coefficients side by side. In particular, they show the standard errors of model coefficients, which set confidence intervals which directly inform the significance of differences between coefficient values from different modelling and data treatment options (para. 30), and the significance of individual coefficients (para. D43).

Different data treatments underly results in the paper and the published code

The published code imports and processes data, fits regression models and collates model coefficients for the full date range 2013-2022.

Data treatments include the deflation of nominal opex and capex amounts into 2019 dollars before model fitting. This is done using the combined index used for opex deflation in DPP3 (60/40 weighted LCI/PPI, para D56).

An outlier exclusion method is applied to iteratively fit a model then exclude data points that fail 3 (or 4) out of 4 outlier tests, until no points fail more than 2 of these 4 tests (para D29).

In comparison, the models shown in the paper tables D6 and D7 reflect other data treatment options, as we explored the modelling aspects we need to determine:

- The date range is noted as 2016-2022. This goes to the option of a sliding date-range window, rather than extending the date range. This trades off increasing the number of data points with the idea that the most recent past is more indicative of the future.
- The approach to removing outliers is the same for the network and non-network opex models (in Tables D6 and D7, respectively), and excludes a fixed set of data points. This is more transparent than the iterative method which by design may remove different points for different model fits and date ranges.

- The capex deflator for Tables D6 and D7 is the Capital Goods Price Index (GCPI) previously used for capex deflation in DPP3.

For Table D6, network opex models, the data points excluded match those excluded by the iterative method in the published code, and the results for network cost models 1,2,3 match the published code when data are filtered to 2016-2022. However, results from the model with a capex term (network opex \sim capex + lines + icp) are slightly different from the published code when data filtered to 2016-2022, due to the different capex deflation.

For Table D7, non-network models, fewer data points are excluded than by the iterative method applied in the published code, and the results for non-network opex models 1,2,3 do not match the comparable models from published code when data are filtered to 2016-2022. For the non-network opex model with a capex term (non-network opex \sim capex + lines + icp) there is also the different capex treatment.

The accompanying script generates tables with the published code results and allows exploration of the impact of changing date range and capex deflation on model fits

The R script *ReEstimates_with_tables.R* attached is a modified version of the published file *ReEstimates.R* and it can be run in the same way.

We've added output tables formatted like the paper tables D6 and D7 showing results from the models in the paper using the data treatment in the original published code:

- date range 2013-2022;
- iterative method of outlier exclusion; and
- existing capex deflation (same for opex and capex).

At the top of the file, the date range can be changed, and the capex deflation series retained or amended to assess the impact on these changes on model fits.

As with the published code, there are also models fit for: other combinations of independent variables, no outlier exclusion, granular opex categories, and robust standard errors. The outlier method underlying the issues paper tables D6 and D7 is not included. All results are collated in the output.csv file.

For any follow ups, please contact us at infrastructure.regulation@comcom.govt.nz

Output tables from attached R script with model options and results from published code

comparison of log-log Network Opex Models 2013-2022; capex deflation: original; outliers: iterative removal

Dependent variable:				
	lines+icp (1)	peak+lines+icp (2)	network delivery+lines+icp (3)	capex+lines+icp (4)
peak		-0.027 (0.051)		
delivery			0.065 (0.066)	
ExpenditureOnAssets				0.008 (0.037)
lines	0.508*** (0.033)	0.516*** (0.033)	0.501*** (0.035)	0.497*** (0.037)
ICP	0.443*** (0.027)	0.466*** (0.060)	0.381*** (0.063)	0.452*** (0.035)
Constant	-0.197 (0.157)	-0.369 (0.394)	0.098 (0.319)	-0.305* (0.155)
Observations	287	285	285	279
R2	0.919	0.920	0.921	0.929
Adjusted R2	0.918	0.920	0.920	0.928
Residual Std. Error	0.276 (df = 284)	0.274 (df = 281)	0.274 (df = 281)	0.261 (df = 275)
F Statistic	1,606.554*** (df = 2; 284)	1,084.458*** (df = 3; 281)	1,087.336*** (df = 3; 281)	1,195.368*** (df = 3; 275)
Note: *p<0.1; **p<0.05; ***p<0.01				

comparison of log-log Non-network Opex Models 2013-2022; capex deflation: original; outliers: iterative removal

Dependent variable:				
	lines+icp (1)	peak+lines+icp (2)	nonnetwork delivery+lines+icp (3)	capex+lines+icp (4)
peak		0.107* (0.059)		
delivery			-0.016 (0.077)	
ExpenditureOnAssets				0.229*** (0.037)
lines	0.279*** (0.036)	0.260*** (0.036)	0.278*** (0.037)	0.163*** (0.039)
ICP	0.599*** (0.030)	0.492*** (0.070)	0.619*** (0.075)	0.450*** (0.036)
Constant	0.393** (0.171)	1.152** (0.451)	0.279 (0.368)	0.690*** (0.167)
Observations	283	279	280	284
R2	0.901	0.903	0.904	0.913
Adjusted R2	0.900	0.902	0.903	0.912
Residual Std. Error	0.300 (df = 280)	0.300 (df = 275)	0.301 (df = 276)	0.283 (df = 280)
F Statistic	1,273.155*** (df = 2; 280)	849.610*** (df = 3; 275)	864.078*** (df = 3; 276)	981.526*** (df = 3; 280)
Note: *p<0.1; **p<0.05; ***p<0.01				