Quantonomics

QUANTITATIVE ECONOMICS

Memorandum

Date:	11/11/2025
From:	Michael Cunningham, Joe Hirschberg, Alice Giovani
To:	Adam Rapoport, Sasha Jergic, Su Wu, Cameron Martin, Arwin Lee
Subject:	Nonreliability Output Index Weights ABR25 – Supplementary Analysis

Contents

1 Introduction

The Quantonomics memo 'Nonreliability Output Index Weights ABR25' of 13 June 2025 updated the output weights for DNSPs and TNSPs, previously last estimated in 2020. The previous weights were based on data from 2006-2018.[1] The updated output weights followed, and were in response to, an independent review by the Centre for Efficiency and Productivity Analysis (CEPA) at the University of Queensland (Peyrache, 2024). They are based on the same methodology as used previously by Economic Insights and use data for 2006-2023.

The updated output weights memo received critical comment from several stakeholders, including AusNet, Evoenergy, SA Power Networks (SAPN) and Ausgrid. Favourable comments were made by Jemena, Ergon/Energex and TasNetworks. A major concern put forward in the submissions is that the updated weights differ materially from those estimated in 2020.

The purpose of this memo is to identify the factors driving the substantial changes in the updated set of output weights and to address stakeholders' comments. Section 2 of this supplementary analysis investigates reasons for the substantial changes in the index weights between the 2020 and 2025 studies.[2] Subsequent sections investigate issues that form the main criticisms of stakeholders. Frontier Economics, commissioned by SAPN and Evoenergy, is critical, among other things, of the starting values used for parameter estimation in the nonlinear least squares (NLS) estimation. This issue is investigated in Section 3.

SAPN and Frontier Economics are also critical of the reliability of the econometric Leontief input demand models which are used to calculate the output weights in the Economic Insights method. These issues are addressed in section 4 through an investigation of regression residuals and goodness-of-fit statistics. Some of the stakeholders' methodological criticisms and other comments are discussed in Section 5. SAPN suggested that alternative ways of averaging the output cost shares across DNSPs should be considered. One alternative method is investigated in Section 5. Lastly, section 6 summarises the conclusions.

2 Results using Different Data Samples

This section examines the reasons for the changes in estimated output weights compared to those published in 2020. Given the methodology for calculating the estimates has been maintained, the reasons for the changes may be:

(a) the inclusion of an additional five years of data (2019-2023) and

[1] DNSP output cost weights were first updated in 2018 using data from 2006–2017, after remaining fixed since the start of economic benchmarking in 2014. Although the intention was to hold these weights constant for five years, a further update was undertaken in 2020 (using 2006–2018 data) to correct an identified time-trend error (Economic Insights, 2020: 4).

[2] This analysis of the changes output weights is limited to the Economic Insights method and does not include the methods for cross-checking proposed by CEPA and implemented in the June 2025 Quantonomics memo.

(b) revisions to data, including

- a redefinition of opex input to include capitalised corporate overheads (CCO) and associated changes to the annual user cost (AUC) of capital inputs
- changes in the AUC calculations to address issues with inflation.

The results presented in this section all use the Economic Insights method, applied using Stata's **nl** command for nonlinear regression, and using starting values of 0.001 for all parameters. Unless otherwise stated, the ABR25 dataset is generally used in this section and restricted as needed.[3] The updated non-reliability output weights in the June 2025 memo are based on the ABR25 dataset restricted to the 2006-2023 period.

2.1 Effects of Changes in Data Definitions

Table 2.1 presents a comparison of the output cost shares estimated with the 2006-2018 data sample using the ABR19 dataset (covering 2006-2018), with which they were originally estimated, and using the ABR25 dataset (for years 2006-2018 only). This comparison shows the effects of data revisions, such as the revised treatment of CCOs in opex and AUC for DNSPs, and the changes in AUC calculation to address issues with inflation.

The changes in data definitions have had the effects of:

- Substantially increasing the weight attributed to RMD (by 5.63 percentage points), and
- Reducing the weights of other outputs, especially Energy Throughput.

Table 2.1 DNSP Output Cost-Share Weights with Different Datasets (2006-2018)

	*ABR19**	*ABR25*	*Diff (p.p.).*
Energy Throughput	8.58%	5.31%	-3.27%
RMD	33.76%	39.39%	5.63%
Customers	18.52%	17.83%	-0.69%
Circuit Length	39.14%	37.47%	-1.67%
Total	100.00%	100.00%	0.00%

Note: *From Economic Insights (2020:4)

2.2 Trends and volatility of Cost Share Weights

Sequentially adding a single year to the sample produces a series of output cost share estimates using samples from 2006-2018 to 2006-2024 using the ABR25 dataset. These are presented in Table 2.2.

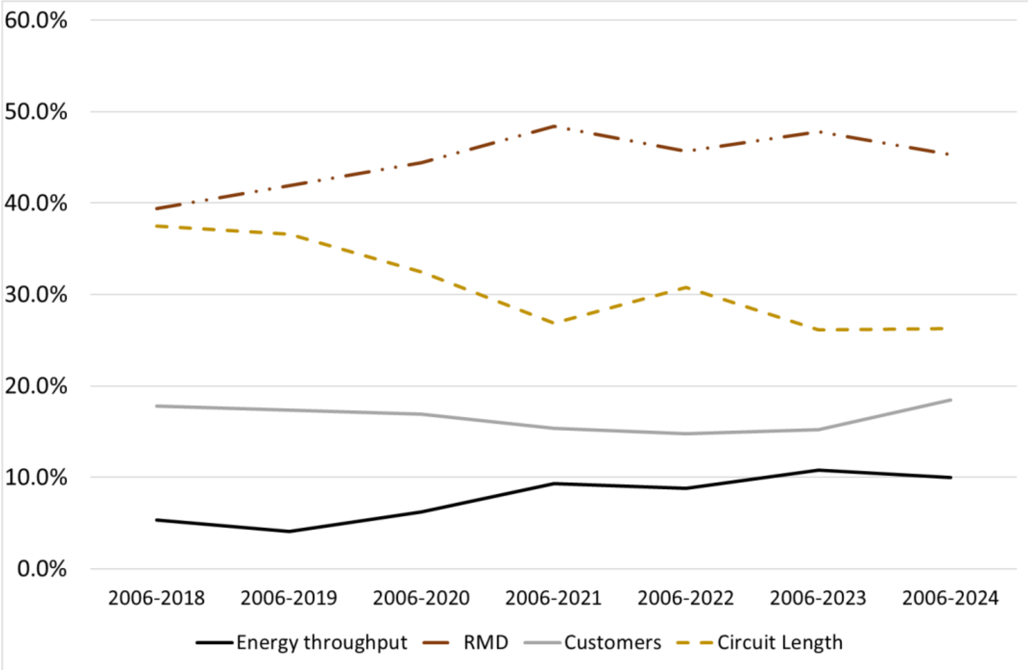
[3] We refer to the dataset used in the 2025 Annual Benchmarking Report (ABR) as ABR25, which covers the period 2006–2024. Similarly, ABR19 refers to the dataset used in the 2019 ABR, covering 2006–2018. The distinction between these datasets is important, as each ABR release incorporates different data revisions. By using the most recent dataset (i.e. ABR25), we ensure that all available data corrections and revisions are reflected in the analysis.

Table 2.2 DNSP Output Cost-Share Weights with Different Data Samples

	2006-2018	*2006-2019*	*2006-2020*	*2006-2021*	*2006-2022*	*2006-2023*	*2006-2024*
Energy Throughput	5.31%	4.10%	6.19%	9.35%	8.77%	10.79%	9.96%
RMD	39.39%	41.89%	44.41%	48.43%	45.67%	47.83%	45.31%
Customers	17.83%	17.39%	16.92%	15.37%	14.81%	15.23%	18.44%
Circuit Length	37.47%	36.62%	32.48%	26.85%	30.76%	26.15%	26.29%
Total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Note: Results were obtained using the ABR25 dataset, restricted to the relevant shown.

Figure 2.1 DNSP Output Cost-Share Weights with Different Data Samples



The results shed light on:

- whether the addition of any specific years results in substantial changes to the estimates;
- trends in output cost share weights over successive samples; and
- the degree of volatility of output cost share weights.

Figure 2.1 also presents the patterns of the output share weights as the sample is incrementally expanded.

Table 2.2 and Figure 2.1 show that whilst there is some volatility in the output weights for DNSPs when estimated over the different data samples, of more importance are the shifts, *especially when the sample was expanded to include 2020 and 2021*, towards higher cost shares for Energy Throughput and RMD and a lower cost share for Circuit Length. From 2006-2021 to 2006-2024, the cost share estimates have been reasonably stable.

2.3 Special factors affecting 2020 and 2021

This section further examines the effects of additional data on output weights. The two main questions relate to (a) structural changes in the input demand functions as the sample has increased; and (b) the effects of changes in input prices on the cost shares of the different inputs, combined with the differences between inputs in the relative importance of the outputs as cost drivers.

To investigate structural changes, Table 2.3 aggregates the estimated coefficients of all the opex input demand functions of the 13 DNSPs for each sample period. Specifically, the squared coefficients applying to each output and the coefficient on the time trend have each been summed across DNSPs, to illustrate, at the industry level, the trends of the coefficients between data samples. Similarly, Tables 2.4 to 2.6 show the aggregated coefficients relating to the overhead lines ("OH Lines"), underground cables ("UG cables") and transformers, respectively.

Table 2.3 shows a weakening of the relationship between opex and circuit length in the 2006-2020 and 2006-2021 data samples. The aggregate squared coefficient applying to circuit length falls from 21.04 in the 2006-2019 data sample, to 12.23 in the 2006-2020 and 0.69 in the 2006-2021. Although there is a partial reversal in the 2006-2022 sample, in the following two data samples, the relationship between opex and circuit length is similar to the 2006-2021 sample. This change may be related to the efficiencies achieved in opex. Commensurately, there have been trends toward greater effects of Energy and RMD.

Table 2.3 Aggregate Opex Function Coefficients

	2006-2018	2006-2019	2006-2020	2006-2021	2006-2022	2006-2023	2006-2024
$\hat{a}_{11}^2$ (Energy)	27.80	24.86	36.22	49.19	40.27	45.01	54.31
$\hat{a}_{12}^2$ (RMD)	334.78	375.09	384.89	445.23	423.41	461.41	405.47
$\hat{a}_{13}^2$ (Cust.)	0.47	0.46	0.49	0.51	0.53	0.61	0.61
$\hat{a}_{14}^2$ (CircLen.)	23.22	21.04	12.23	0.69	8.94	0.58	0.60
$\hat{b}_1$	0.18	0.13	0.09	0.06	0.01	0.03	0.04

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years shown.

Tables 2.4 to 2.6 show the aggregated function for the three capital inputs, and all show some trend toward greater importance of RMD. In Table 2.6, Transformer demand is also more strongly affected by Energy Throughput in the 2006-2021 to 2006-2023 data samples, but this effect is no longer present in the 2006-2024 sample. Aside from these effects, the capital inputs do not appear to show major structural shifts and differ in this regard from the opex input demand function.

Table 2.4 Aggregate OH Lines Function Coefficients

	2006-2018	2006-2019	2006-2020	2006-2021	2006-2022	2006-2023	2006-2024
$\hat{a}_{21}^2$ (Energy)	10.88	9.38	11.38	9.56	11.28	10.40	9.61
$\hat{a}_{22}^2$ (RMD)	167.25	154.68	127.02	177.82	203.96	227.20	229.64
$\hat{a}_{23}^2$ (Cust.)	0.57	0.57	0.58	0.56	0.57	0.56	0.55
$\hat{a}_{24}^2$ (CircLen.)	59.66	61.06	65.07	54.97	51.15	50.18	49.75
$\hat{b}_2$	-0.03	-0.05	-0.06	-0.04	-0.04	-0.04	-0.04

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years shown.

Table 2.5 Aggregate UG Cables Function Coefficients

	2006-2018	2006-2019	2006-2020	2006-2021	2006-2022	2006-2023	2006-2024
$\hat{a}_{31}^2$ (Energy)	0.19	0.00	0.17	0.27	0.62	0.53	1.06
$\hat{a}_{32}^2$ (RMD)	22.91	23.45	25.65	26.55	26.58	27.43	36.55
$\hat{a}_{33}^2$ (Cust.)	0.08	0.09	0.07	0.07	0.07	0.07	0.06
$\hat{a}_{34}^2$ (CircLen.)	9.39	9.51	9.11	8.86	8.71	9.05	7.39
$\hat{b}_3$	0.45	0.44	0.44	0.43	0.42	0.42	0.42

Note: Results were obtained using the ABR25 dataset, restricted to the relevant shown.

Table 2.6 Aggregate Transformer Function Coefficients

	2006-2018	2006-2019	2006-2020	2006-2021	2006-2022	2006-2023	2006-2024
$\hat{a}_{41}^2$ (Energy)	1.27	1.11	0.77	2.15	2.00	1.69	0.47
$\hat{a}_{42}^2$ (RMD)	10.59	12.45	14.49	14.57	16.28	16.79	16.14
$\hat{a}_{43}^2$ (Cust.)	0.06	0.06	0.05	0.04	0.03	0.04	0.06
$\hat{a}_{44}^2$ (CircLen.)	2.29	2.24	2.50	2.31	2.31	2.44	2.29
$\hat{b}_4$	0.24	0.23	0.21	0.21	0.20	0.18	0.17

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years shown.

Another data change relevant to 2020 and 2021 is the substantial decrease in capital input prices in those two years, which increases the relative importance of the opex input in total cost. Table 2.7 shows the trends in the relative importance of each input in total cost between the data samples. The cost share of opex began to increase in the 2006-2020 data sample and increased further in each sample up to 2006-2023.[4] Among the capital inputs, UG Cables had the largest decline.

Table 2.8 shows the shares of opex cost estimated to be attributable to each output in each data sample. Tables 2.9 to 2.11 show the shares of each capital input cost estimated to be attributable to each output in each data sample. Table 2.8 shows the majority of Opex is driven by RMD, so the increased cost share of opex caused RMD to increase in overall importance. Table 2.8 also highlights the effects of the structural changes in the opex input demand function previously noted. A substantial decline in the proportion of Opex attributed to the circuit length output, an increase in the proportion attributed to energy, and an increase in the proportion of cost attributed to RMD between the 2006-2018 and 2006-2023 samples (although this decreased in 2006-2024).

Table 2.7 Input Total Cost Shares (%)

Data sample	*Opex*	*OH Lines*	*UG Cables*	*Transformers*	*Total*
2006-2018	41.9%	18.5%	13.0%	26.6%	100.0%
2006-2019	41.9%	18.5%	13.0%	26.6%	100.0%
2006-2020	42.0%	18.5%	12.9%	26.5%	100.0%
2006-2021	42.3%	18.4%	12.9%	26.5%	100.0%
2006-2022	42.4%	18.4%	12.8%	26.4%	100.0%
2006-2023	42.5%	18.4%	12.7%	26.4%	100.0%
2006-2024	42.3%	18.5%	12.7%	26.5%	100.0%

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years shown.

Table 2.8 Opex Cost: Per cent Attributable to Outputs

Data sample	*Energy*	*RMD*	*Cust*	*CircLen*	*Total*
2006-2018	8.4%	56.9%	14.9%	19.8%	100.0%
2006-2019	7.0%	60.5%	14.8%	17.7%	100.0%
2006-2020	11.3%	63.9%	14.9%	9.9%	100.0%
2006-2021	11.7%	70.6%	15.2%	2.5%	100.0%
2006-2022	10.4%	61.1%	15.4%	13.1%	100.0%
2006-2023	15.9%	65.6%	16.4%	2.1%	100.0%
2006-2024	21.0%	58.9%	16.4%	3.7%	100.0%

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years shown.

Tables 2.9 and 2.10 show that the major cost driver of OH Lines and UG Cables is circuit length, which is logical given that extensions to the network will need additional OH Line and

[4] Table 2.7 shows incremental impact of adding one more year on the input cost shares for the entire sample period examined. The implied change in the input cost share for the additional one year is more significant.

UG Cable inputs. However, with additional data, circuit length has decreased in relative importance as a cost driver for these two inputs. This may be due to an increasing proportion of the growth in OH Lines and UG Cables coming from line/cable capacity re-rating.

Table 2.9 OH Lines Cost: Percent Attributable to Outputs

Data sample	*Energy*	*RMD*	*Cust*	*CircLen*	*Total*
2006-2018	2.8%	12.9%	10.9%	73.4%	100.0%
2006-2019	2.2%	12.1%	10.8%	74.9%	100.0%
2006-2020	2.8%	11.5%	11.0%	74.7%	100.0%
2006-2021	2.2%	15.3%	11.0%	71.5%	100.0%
2006-2022	2.7%	16.9%	11.0%	69.4%	100.0%
2006-2023	2.4%	17.7%	10.7%	69.2%	100.0%
2006-2024	2.2%	17.9%	10.9%	69.0%	100.0%

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years shown.

Table 2.10 UG Cables Cost: Percent Attributable to Outputs

Data sample	*Energy*	*RMD*	*Cust*	*CircLen*	*Total*
2006-2018	0.9%	28.3%	10.0%	60.7%	100.0%
2006-2019	0.0%	29.1%	10.2%	60.7%	100.0%
2006-2020	0.8%	32.8%	8.9%	57.6%	100.0%
2006-2021	1.2%	34.5%	13.3%	51.0%	100.0%
2006-2022	2.9%	34.2%	15.3%	47.5%	100.0%
2006-2023	2.6%	34.8%	15.1%	47.5%	100.0%
2006-2024	2.9%	41.6%	15.0%	40.5%	100.0%

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years shown.

Table 2.11 Transformers Cost: Percent Attributable to Outputs

Data sample	*Energy*	*RMD*	*Cust*	*CircLen*	*Total*
2006-2018	4.3%	35.4%	31.2%	29.0%	100.0%
2006-2019	2.8%	39.8%	29.5%	27.8%	100.0%
2006-2020	3.2%	42.4%	28.3%	26.2%	100.0%
2006-2021	14.5%	42.9%	19.7%	22.9%	100.0%
2006-2022	13.2%	46.5%	16.3%	24.0%	100.0%
2006-2023	12.5%	47.1%	16.8%	23.7%	100.0%
2006-2024	1.2%	44.5%	28.7%	25.6%	100.0%

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years shown.

In the 2006-2018 sample, Transformer costs were attributable in roughly equal proportions to RMD, Customer Numbers and Circuit Length. However, in 2006-2023, a much larger proportion of Transformer costs was attributable to RMD and Energy, and smaller proportions were attributable to Customer numbers and Circuit Length. These changes have been partly reversed in the 2006-2024 data sample, especially the cost share attributable to

Energy. The 2006-2024 data sample appears to produce a more consistent attribution of Transformer costs than the 2006-2023 sample.

2.4 Summary observations

Table 2.12 summarises the effects of the data revisions and the additional data, comparing the 2006-2018 dataset with the 2006-2023 dataset. The second column reproduces the last column of Table 2.1. The third column represents the estimated effect of additional data. It is obtained by subtracting the 2nd column of Table 2.2 (the output cost shares using the 2006-2018 sample) from the 2nd-last column of Table 2.2 (the output cost shares estimated using the 2006-2023 sample. The last column of Table 2.12 shows the combined effects. This Table shows:

- The large increase in the cost share of RMD is a combination of increases due to data revisions and to additional data, both considerably raising its weight;
- The modest increase in the cost share of Energy Throughput is due to largely offsetting effects from data revisions and additional data;
- The large decrease in the cost share of Circuit Length is mostly due to the additional data, which has tended to reduce the estimated influence of this output on input demands.

Table 2.12 Decomposition of DNSP Output Cost-Share Changes between 2006-2018 and 2006-2023*

	Diff. due to data revisions	*Diff. due to additional data*	*Total Change*
Energy Throughput	-3.27%	5.48%	2.21%
RMD	5.63%	8.44%	14.07%
Customers	-0.69%	-2.60%	-3.29%
Circuit Length	-1.67%	-11.32%	-12.99%
Total	0.00%	0.00%	0.00%

Note: * Using ABR25 dataset.

The effects of additional data were mostly associated with the addition of 2020 and 2021 data. They appear to be driven by the following main factors:

- The weakening of the relationship between opex and circuit length in the 2006-2020 and 2006-2021 data samples, perhaps related to the efficiencies that have been achieved in opex.
- A relatively steady trend towards reduced importance of circuit length as a driver of OH Lines and UG Cables costs, perhaps due to changing (increasing) line capacities.
- A relatively steady trend toward greater attribution of costs to RMD for all inputs.
- A large reduction in the prices of capital inputs in 2020 and 2021, which increased the share of opex in total costs. Since a relatively high share of Opex is attributed to RMD

and a relatively low share is attributed to circuit length, this accentuated the shift of overall cost share away from Circuit Length towards RMD.

3 Alternative Starting Values

3.1 Frontier Economics' Analysis

Evoenergy's submission states:

> Frontier Economics has identified that Quantonomics did not estimate the Leontief model correctly in six instances. Quantonomics uses the nl package in Stata to estimate each model. … Using a different approach (maximum likelihood estimation), Frontier Economics identified six models that had materially different coefficient estimates to the results presented by Quantonomics. When those coefficient estimates were fed into the nl Stata package as starting values, the solver identified a different set of estimated coefficients than results presented by Quantonomics, with an improved model fit (i.e., lower sum of squared residuals). This means that the Quantonomics estimates were not valid estimates as they did not minimise the sum of squared residuals. (Evoenergy, 2025)

The Quantonomics output weights memo relied only on the method used by Economic Insights and the key methods proposed by CEPA. Frontier Economics' approach of using maximum likelihood (ML) estimation to obtain alternative starting values for applying the Economic Insights method is of interest as a potential improvement (albeit adding an extra step in the process of the NLS estimation, to first derive the alternative starting values).

It can be taken as given that a method employed for some DNSPs must also be used for all DNSPs. It would not be satisfactory to use different methods for different DNSPs in some *ad hoc* manner. A method that cannot be consistently applied to *reliably* produce results for all DNSPs cannot be used to produce a reliable set of output weights, because some of the constituent models are unreliable. Furthermore, since a consistent method is needed to produce the average output weights, a method that performs better for 6 of the 52 input demand models does not necessarily consistently produce superior results over all models and is therefore not necessarily a better method or even a feasible method.

In this section, we estimate the input demand models by first estimating the models using Stata's **ml** functionality, which allows the user to specify the log likelihood function to be minimised. The convergence properties of this estimator are discussed. The resulting estimated coefficients are then used as starting values in the nonlinear least squares (NLS) estimation of input demand functions. For each regression (for each DNSP and input), the resulting sum of squared residuals is compared to those obtained using the Economic Insights method.

3.1.1 Convergence of ML models

Table 3.1 shows the convergence results of ML models using the 2006-2023 and 2006-2024 samples. For these ML models, we use the standard 0.001 as starting values, in line with Frontier Economics' approach. There are 52 models to be estimated in total. In both samples, two of the ML models failed to converge. This means that the method of using ML estimates to obtain starting values cannot be applied consistently because in some cases they do not converge. In the cases where the ML models did not successfully converge, we have used parameter starting values of 0.001 in a single step for the NLS estimation.[5]

Table 3.1 Convergence of ML Models

Sample period	*Models failing to achieve convergence*
2006-2023	• ESS UG/Cables • ESS Transformers
2006-2024	• UED UG/Cables • UED Transformers

3.1.2 Results of Using ML Estimates as Starting Values

The purpose is to test whether the root mean square error (RMSE) of each input demand model is lower (or no higher) when the ML estimates are used as starting values than when the starting values used by Economic Insights are used.[6] Tables 3.2 and 3.3 present comparisons of the RMSEs for each model using the standard starting values compared to the RMSEs obtained when the ML estimates are used as starting values. Table 3.2 uses the 2006-2024 sample period, and Table 3.3 uses the 2006-2023 period. In both sample periods:[7]

- 15 of the models have a lower RMSE when the ML starting values are used
- 2 of the models have a higher RMSE when the ML starting values are used. In the 2006-2024 sample, AGD UG cables and TND transformers. In the 2006-2024 sample, PCR real opex and TND transformers.
- For the remaining 35 models, the RMSE is unchanged.

On balance, the use of the ML starting values tends to produce some improvement in lowering the RMSE. On average over *both* sample periods (ie, 104 models), the percentage reduction in RMSE is 3.3 per cent.

[5] For the output-related parameters a_{ij} (for input demand i and output j), which are squared in the Economic Insights specification, we have used the absolute values of the first-round estimates as starting values for the second-round estimation.

[6] Comparing RMSE values is one method of comparing the goodness-of-fit of econometric specifications (Baum, 2006: 79). Unlike R^2, the RMSE is not bounded between 0 and 1 and expresses the average distance between predicted and actual values in absolute terms, whereas R^2 indicates the proportion of variance in the dependent variable explained by the model.

[7] A threshold of 1 per cent is used when we refer to an RMSE being higher or lower than another.

Table 3.2 RMSE by Input Demand Model, 2006-2024 data sample

	Standard starting Values				ML starting values			
DNSP	*x1*	*x2*	*x3*	*x4*	*x1*	*x2*	*x3*	*x4*
1. EVO	8,986	2,209	747	39	8,811	2,209	720	39
2. AGD	58,507	6,948	4,255	1,528	58,507	6,514	4,255	1,528
3. CIT	4,198	140	239	67	3,990	140	239	67
4. END	17,649	6,201	1,788	402	17,649	6,201	1,788	402
5. ENX	20,104	3,132	4,196	679	20,104	3,132	4,196	655
6. ERG	31,743	13,280	805	426	31,743	13,280	781	413
7. ESS	49,316	8,120	1,876	743	49,316	8,120	1,815	686
8. JEN	8,777	1,530	293	106	8,337	870	293	106
9. PCR	10,993	1,314	648	86	11,301	1,314	648	83
10. SAP	15,068	1,651	695	288	14,941	1,651	695	288
11. AND	13,479	2,625	389	178	13,479	2,625	389	173
12. TND	7,740	662	160	84	7,081	662	160	86
13. UED	8,262	2,122	134	104	8,262	2,054	134	104

Table 3.3 RMSE by Input Demand Model, 2006-2023 data sample

	Standard starting Values				ML starting values			
DNSP	*x1*	*x2*	*x3*	*x4*	*x1*	*x2*	*x3*	*x4*
1. EVO	9,075	2,356	341	40	8,979	2,274	341	40
2. AGD	62,310	6,708	3,971	1,346	60,264	6,708	4,066	1,346
3. CIT	4,101	140	246	68	4,101	140	229	68
4. END	18,853	6,400	1,653	402	18,192	6,400	1,653	402
5. ENX	19,999	5,992	3,150	518	19,999	3,108	3,150	518
6. ERG	25,070	43,575	781	414	25,070	13,632	781	414
7. ESS	49,582	8,120	1,874	674	49,582	8,120	1,874	674
8. JEN	8,362	859	253	89	8,362	859	244	86
9. PCR	11,802	1,357	716	83	11,802	1,357	665	83
10. SAP	15,412	1,708	729	308	15,057	1,708	729	294
11. AND	14,222	24,689	345	175	13,770	2,349	345	175
12. TND	7,300	684	165	85	7,300	684	165	88
13. UED	8,395	2,172	138	107	8,395	2,098	138	107

Note: Results were obtained using the ABR25 dataset, restricted to 2006-2023.

3.1.3 Implications for Output Cost Shares

In Table 3.4, the output cost shares estimated using the models that use the ML starting values are compared to those estimated using the standard starting values. These comparisons show that, compared to the standard starting values, the ML starting values produce:

- lower cost shares for Energy Throughput, particularly in the 2006-2023 data sample, with a much smaller effect in the 2006-2024 sample;
- higher cost shares for RMD by approximately 4 percentage points in both samples;
- the customer numbers cost share is similar in the 2006-2023 sample, but lower in the 2006-2024 sample; and

- cost shares for Circuit Length are similar in both sample periods.

Table 3.4 Comparison of Output Cost Shares with ML Starting Values

	Standard starting values		*ML starting values*	
	2006-2023	*2006-2024*	*2006-2023*	*2006-2024*
Energy Throughput	10.79%	9.96%	6.78%	8.08%
RMD	47.83%	45.31%	52.06%	49.11%
Customers	15.23%	18.44%	15.12%	15.33%
Circuit Length	26.15%	26.29%	26.04%	27.48%
Total	100.00%	100.00%	100.00%	100.00%

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years as shown.

As shown in Table 2.12, the main changes from the weights estimated by Economic Insights in 2020 are the large increase in the cost share of RMD and the decrease in the cost share of Circuit Length. Table 3.5 shows that Frontier Economics' method of using ML starting values does not substantially alter the main changes in weights. The changes from the previous weights remains quite significant and move broadly in the same direction as those observed with our original estimates (the only exception being energy throughput).

Table 3.5 Difference from Previous Weights

	Standard starting values		*ML starting values*	
	2006-2023	*2006-2024*	*2006-2023*	*2006-2024*
Energy Throughput	2.21%	1.38%	-1.80%	-0.50%
RMD	14.07%	11.55%	18.30%	15.35%
Customers	-3.29%	-0.08%	-3.40%	-3.19%
Circuit Length	-12.99%	-12.85%	-13.10%	-11.66%
Total	0.00%	0.00%	0.00%	0.00%

3.2 Preceding Sample Estimates as Starting Values

Another approach to starting values is to use previously estimated parameters as starting values. This is the method used by CEPA. This method can be extended to sequentially estimate output weights using parameters previously estimated with the previous, smaller data sample.[8] Table 3.6 presents output cost share results obtained using each sample period from 2006-2018 to 2006-2024. Figure 3.1 depicts the patterns of output weights obtained. An important feature is the large shifts to greater RMD weight and smaller Circuit Length weight.

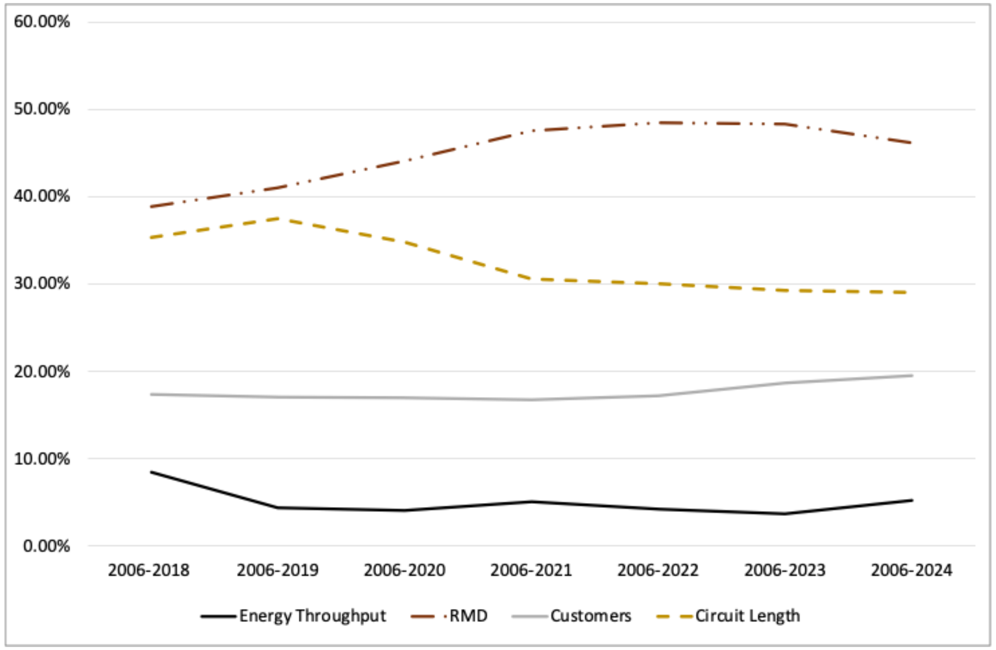
[8] For the 2006-2018 data sample, the Economic Insights parameter estimates are used as starting values. The models applied to the 2006-2019 dataset use the coefficients estimated with the 2006-2018 data, and so on, by adding one year at a time, and using the estimated coefficients of the preceding sample as starting values for the next (more specifically, the absolute values of the squared parameters).

Table 3.6 DNSP Output Cost-Share Weights with Different Data Samples—Starting Values from Preceding Results

	2006-2018	*2006-2019*	*2006-2020*	*2006-2021*	*2006-2022*	*2006-2023*	*2006-2024*
Energy Throughput	8.44%	4.39%	4.05%	5.05%	4.22%	3.68%	5.25%
RMD	38.89%	41.05%	44.10%	47.61%	48.50%	48.39%	46.18%
Customers	17.34%	17.06%	16.99%	16.77%	17.21%	18.67%	19.56%
Circuit Length	35.33%	37.50%	34.85%	30.57%	30.07%	29.26%	29.02%
Total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years as shown.

Figure 3.1 DNSP Output Cost-Share Weights with Different Data Samples—Starting Values from Preceding Results



Tables 3.7 and 3.8 present comparisons of the RMSEs for each model using the standard starting values compared to the RMSEs obtained when the preceding estimates are used as starting values. Table 3.7 uses the 2006-2024 sample period, and Table 3.8 uses the 2006-2023 period.

The results in Table 3.6 for the 2006-2023 and 2006-2024 data samples can be compared to those using standard starting values. This comparison is shown in Table 3.9.

Table 3.7 RMSE by Input Demand Model, 2006-2024 data sample

	Standard starting values				*Preceding results as starting values*			
DNSP	*x1*	*x2*	*x3*	*x4*	*x1*	*x2*	*x3*	*x4*
1. EVO	8,986	2,209	747	39	9,586	2,209	698	39
2. AGD	58,507	6,948	4,255	1,528	58,507	6,514	4,228	1,528
3. CIT	4,198	140	239	67	3,990	140	239	67
4. END	17,649	6,201	1,788	402	17,649	6,201	2,413	402
5. ENX	20,104	3,132	4,196	679	20,104	3,308	4,248	655
6. ERG	31,743	13,280	805	426	31,842	14,372	781	413
7. ESS	49,316	8,120	1,876	743	49,316	8,120	1,815	675
8. JEN	8,777	1,530	293	106	8,636	970	293	103
9. PCR	10,993	1,314	648	86	11,672	1,314	648	86
10. SAP	15,068	1,651	695	288	14,941	1,651	720	288
11. AND	13,479	2,625	389	178	13,479	2,625	401	173
12. TND	7,740	662	160	84	7,360	662	160	84
13. UED	8,262	2,122	134	104	8,262	2,054	145	104

Table 3.8 RMSE by Input Demand Model, 2006-2023 data sample

	Standard starting values				*Preceding results as starting values*			
DNSP	*x1*	*x2*	*x3*	*x4*	*x1*	*x2*	*x3*	*x4*
1. EVO	9,075	2,356	341	40	9,693	2,416	341	42
2. AGD	62,310	6,708	3,971	1,346	60,264	6,708	3,971	1,346
3. CIT	4,101	140	246	68	4,101	140	229	68
4. END	18,853	6,400	1,653	402	18,192	6,400	2,113	402
5. ENX	19,999	5,992	3,150	518	19,999	3,108	3,150	518
6. ERG	25,070	43,575	781	414	25,070	14,795	781	414
7. ESS	49,582	8,120	1,874	674	49,582	8,120	1,874	717
8. JEN	8,362	859	253	89	8,732	938	254	86
9. PCR	11,802	1,357	716	83	11,802	1,357	665	86
10. SAP	15,412	1,708	729	308	15,057	1,708	729	294
11. AND	14,222	24,689	345	175	13,770	2,349	339	175
12. TND	7,300	684	165	85	7,585	684	165	85
13. UED	8,395	2,172	138	107	8,395	2,098	145	107

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years as shown.

Table 3.9 Comparison of Output Cost Shares with Preceding Results as Starting Values

	Standard starting values		*Preceding results as starting values*	
	2006-2023	*2006-2024*	*2006-2023*	*2006-2024*
Energy Throughput	10.79%	9.96%	3.68%	5.25%
RMD	47.83%	45.31%	48.39%	46.18%
Customers	15.23%	18.44%	18.67%	19.56%
Circuit Length	26.15%	26.29%	29.26%	29.02%
Total	100.00%	100.00%	100.00%	100.00%

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years as shown.

The comparisons in Table 3.9 show that the method of using preceding results as starting values produces substantially different results, with a lower weight for Energy Throughput and higher weights for the other outputs, especially Circuit Length.

The comparisons in Tables 3.7 and 3.8 show that:

- In the 2006-2023 sample, 13 of the models have a lower RMSE, and 10 have a higher RMSE, than when the standard starting values are used;
- In the 2006-2024 sample, 14 of the models have a lower RMSE, and 9 have a higher RMSE, than when the standard starting values are used.

The use of the preceding estimates as starting values *lowers* the average RMSE compared to the standard starting values. On average over *both* the 2006-2023 sample and the 2006-2024 samples by 1.8 per cent in sample. When compared to the models using the ML starting values, those estimated using preceding estimates as starting values produce an average RMSE (including averaging over both sample periods) that is 1.7 per cent higher.

Hence, on the basis of average RMSE, the approach of using previously estimated parameters as starting values is arguably superior to the standard starting values and slightly inferior to the use of ML results as starting values. Concerning Frontier Economics' criticism quoted at the beginning of this section, relating to "valid estimates", it is interesting to observe that when the RMSEs of models using preceding results starting values (shown in Tables 3.7 and 3.8) are compared to those using ML estimates as starting values (shown in Tables 3.2 and 3.3):

- In the 2006-2023 sample, 3 of the models that use preceding results starting values have a lower RMSE, and 12 have a higher RMSE, than when the ML starting values are used.
- In the 2006-2024 sample, 4 of the models that use preceding results starting values have a lower RMSE, and 13 have a higher RMSE, than when the ML starting values are used.

Thus, the use of ML starting values does not yield the lowest RMSE for all models. In our view, the average RMSE for all 52 models is a better indicator of the effectiveness of a method of choosing starting values than selecting a small number of models to highlight.

3.3 Conclusions

This section has evaluated Frontier Economics' approach of using parameters estimated using Stata's **ml** routines as starting values for the NLS estimation of the input demand functions used to calculate output cost shares. The following findings are largely invariant to whether the 2006-2023 or the 2006-2024 data samples are used:

- Out of the 52 ML models estimated (for each sample), 2 did not converge. We have used the standard starting values in place of the unreliable ML estimates for those models.
- When the ML-estimated parameters are used as starting values in NLS estimation, 15 of the 52 models have a lower RMSE than when standard starting values are used. However, 2 models have a higher RMSE. Thus, the RMSE is worse for some models, although on average (using both sample periods), the RMSE is improved by 3.3 per cent.
- When previously estimated parameters are used as starting values in NLS estimation, about 12–13 of the 52 models have a higher RMSE than when ML-estimated starting values are used. However, 3–4 models have a lower RMSE than when ML-estimated starting values are used. Thus, several of the models estimated with ML-estimated starting values are worse than models where previously estimated parameters are used as starting values. However, on average, the models using ML-estimated starting values have slightly lower RMSE than models using previously estimated parameters as starting values (by 1.7 per cent on average over both sample periods).

Frontier Economics claimed that "Quantonomics failed to estimate the Leontief model correctly in six instances … the real opex models for Evoenergy and SA Power Networks, the overhead lines model for Ergon Energy and AusNet Distribution, the underground cables model for Ausgrid and the transformers model for TasNetworks Distribution" (SA Power Networks, 2025: 10). This criticism is based on the observation that, when starting values obtained from ML estimates are used, Frontier Economics found 6 models with a lower RMSE than the equivalent models that used the standard starting values. However, as we have seen:

- Some of the ML routines for estimating starting values failed to converge. And when the ML estimates are used as starting values for NLS models, 2 of the models produced a higher RMSE than the equivalent models as starting values. These issues were not noted in the submissions.
- When the preceding estimated coefficients are used as starting values, about 3 models are found with lower RMSE than the equivalent models that used the ML estimates as starting values. Hence, none of the three methods examined produces the lowest RMSE in all 52 input demand models.

These results highlight CEPA's point that NLS estimation produces parameters that yield a local minimum for the sum of squared residuals, but does not reliably find a global minimum (Peyrache, 2024:14).

A useful single metric to use when comparing the three methods considered here for setting starting values is the average RMSE across all 52 input demand models. By this metric, the models that use ML-estimated starting values proposed by Frontier Economics perform the best of the methods tested.

As previously stated, the use of different methods of establishing starting values for different DNSPs and inputs would be *ad hoc*. The method of establishing starting values should be consistently applied to the models for all DNSPs and inputs. We have found that the procedure used by Frontier Economics of first estimating parameters using Stata's ML routine, and then using these as starting values in the NLS estimation produced better results than the standard starting values we used, which were previously used by Economic Insights. If the ML estimation routine fails to converge, the standard starting values (0.001) can be used as a fallback assumption.[9] We recommend that this approach be considered as a potential improvement over the use of the standard starting values in future updates of the output weights.

Frontier Economics also criticises the fact that many of the individual coefficients are zero, or very close to zero, in the four input demand functions estimated for each DNSP. The reason why they consider this to be a deficiency is not explained. The overall aim is to obtain the average effect of output on costs for the industry. The purpose and advantages of estimating the models separately for individual DNSPs were explained by CEPA (2024: 14), namely that it allows a highly flexible estimation procedure. We consider these criticisms by Frontier Economics to be mistaken.

Further, in its Stata program, Frontier Economics describes output coefficients that are nonnegative but close to zero as monotonicity violations. This is a mistake. Monotonicity requires only that the conditional input demand is nondecreasing in any output (Coelli et al., 2005: 25–26). Hence, the output coefficients of the input demands must be nonnegative. An output coefficient equal to or greater than zero is not a monotonicity violation. The squaring of the output parameters ensures that there can be no monotonicity violations in these models.

4 Residuals Analysis and Model Fit

This section examines the residuals from the input demand models for each DNSP to identify any estimation issues or misspecifications. The focus is primarily on the output weights estimated with the 2006-2023 data used in the benchmarking report 2025, although we also examine the residuals when Frontier Economics' suggested use of alternative starting values,

[9] We used the absolute values of the squared parameters as starting values for the second round, although we note that Frontier Economics generally used the estimated values (retaining negative values for starting values where applicable) when examining the six instances with mis-estimation using the standard starting values.

derived from maximum likelihood estimation, is adopted. We also examine the residuals when the 2006-2024 data sample is used.

Appendix A presents the fitted values against the actual values for each of the 52 Leontief estimations, covering the 2006–2023 and 2006–2024 periods and using the ABR25 dataset.

4.1 Identifying Possible Outliers & Residuals Patterns

Outliers can be defined as observations with unusually large residuals, indicating that the dependent variable is not well explained by the predictor variables. To detect outliers, the residuals for each DNSP are first standardised. For each residual e_i (for observation i) the standardised residual is calculated as: $z_i = \frac{e_i - \overline{e_g}}{s_g}$, where $\overline{e_g}$ is the mean residual for DNSP g and s_g is the standard deviation of residuals for DNSP g. This scales residuals relative to the distribution of residuals within each DNSP. Any observation with $|z_i| > 3$ can be considered a potential outlier.

Using this criterion, no observations were identified as potential outliers for any DNSP.

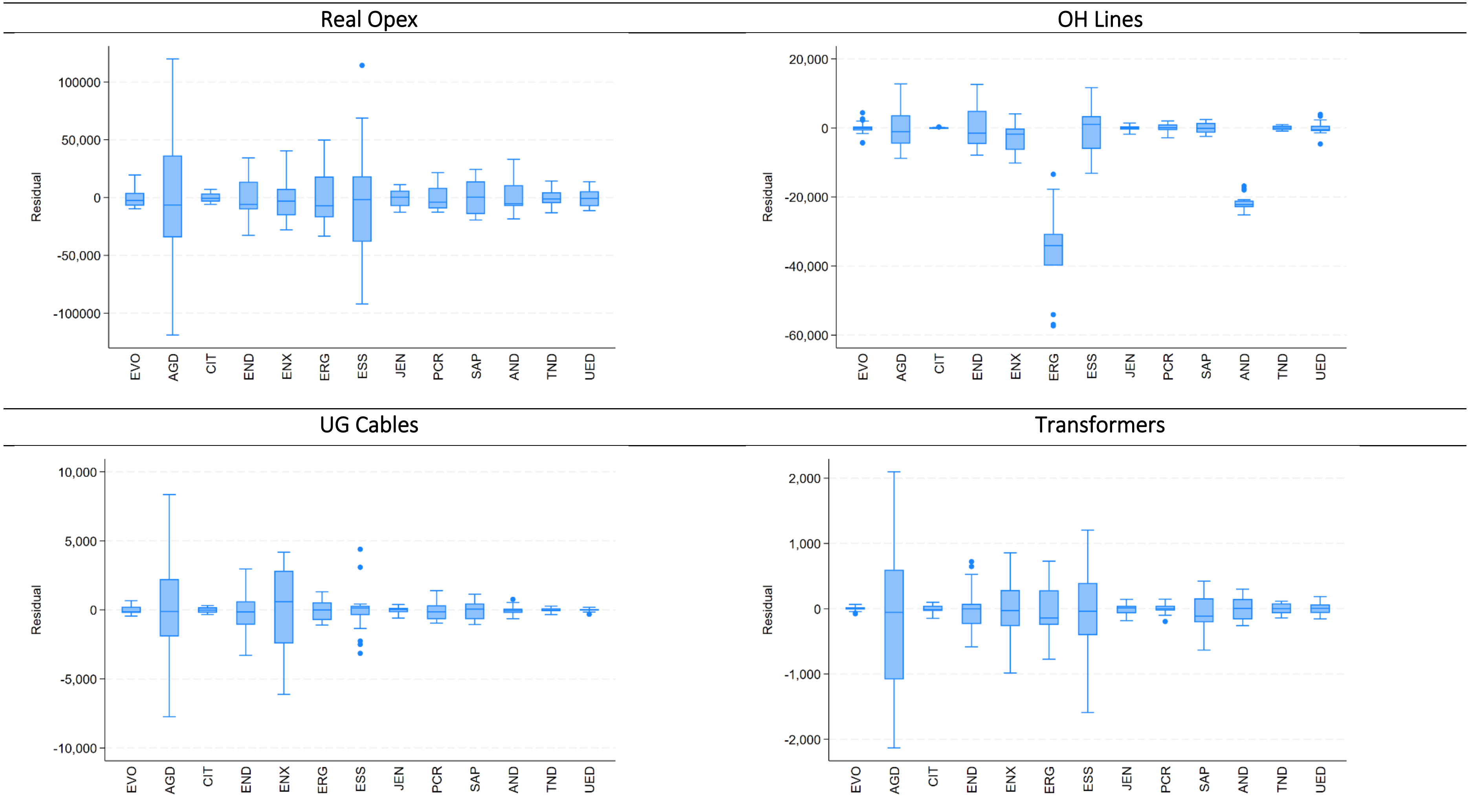
Boxplots are useful for inspecting the distributions of residuals and outliers. A boxplot shows the median (line inside the box), the interquartile range or IQR (the box from the 25th to the 75th percentile), and "whiskers" that extend to the most extreme values within 1.5×IQR from the box. Points plotted beyond the whiskers are flagged as outliers under the standard 1.5×IQR rule, and points beyond 3×IQR are typically considered extreme outliers.

Figure 4.1 presents boxplots of the residuals for each DNSP, across the four inputs: opex, overhead lines, underground cables and transformers, for the 2006-23 period. Note that the vertical axis scale differs across the charts. In practice, the boxplots can be interpreted as follows:

- In a well-specified model, residuals should be clustered around zero.[10] The mean should be close to zero, indicating that, on average, the model's predictions are accurate. A median (the line inside the box) that sits noticeably above or below zero suggests systematic over- or under-prediction for that DNSP.
- Very long whiskers or a large number of outlier points suggest heavy-tailed distributions or the presence of potential outliers.

[10] Because the nonlinear input demand functions do not have intercepts, the residuals need not be centred precisely at zero.

Figure 4.1: Residual Distributions by DNSP (2006-2023)



Note: Results were obtained using the ABR25 dataset, restricted to 2006-2023.

From Figure 4.1, we observe that:

- The residuals for most DNSPs are indeed clustered around zero. However, some notable exceptions arise that may indicate specification issues. The most prominent case is for ERG and AND in OH lines, where all residuals fall well below zero, signalling systematic under-prediction.
- AGD also displays a much wider spread of residuals across all inputs except OH lines. Broader distributions are likewise observed for ESS and ERG in real opex and transformers, and for ENX in UG cables.
- Residuals for real opex exhibit the widest distribution across all DNSPs, followed by OH lines. In contrast, UG cables and transformers show residuals that are more tightly clustered around zero for all DNSPs, suggesting these outputs are generally better explained by the model.

4.2 Residuals in the 2006-2024 dataset

Figure 4.2 presents the boxplots for the 2006–2024 sample. The issue observed in the ERG and AND OH lines models, where residuals were not centred around zero, disappears when the extended dataset is used. However, as discussed in Section 2.2, the estimated output weights remain very similar across the two samples. This suggests that the residuals centring problem in the two models does not materially affect the reliability of the estimated output weights, although we note the improvement with the 2006–2024 sample.

4.3 Goodness-of-Fit

R^2 values of regressions are shown in Table 4.1. The first panel (*EI 2006-2018*) presents R^2 values reported by Economic Insights in 2020. The second panel (*2006–2023*) shows R^2 values from the regressions used to update the output weights, and on which the index analysis in the benchmarking report is based. The third panel (*2006–2024*) are the results when using the full ABR25 dataset. Note that all these estimations use the same set of coefficient starting values (0.001). The main observations on Table 4.1 are:

- *EI 2006–2018*: The average R^2 across the four inputs is slightly higher than in the 2006–2023 dataset. All R^2 values are positive, although a few are quite small, especially for Opex and OH lines.
- *2006–2023*: The R^2 values are mostly positive, with the exceptions of the OH lines input for ERG and AND, which are negative. This reinforces the problematic nature of those two models as previously observed.

  2006–2024: The average R^2 values are similar to those using the 2006-2023 sample, but the issue with the OH Lines input for ERG and AND has disappeared, again showing that the inclusion of 2024 data resolved the computational problems with those models. The R^2 for EVO in the opex estimation is effectively zero but negative.

Figure 4.2: Residual Distributions by DNSP (2006-2024)

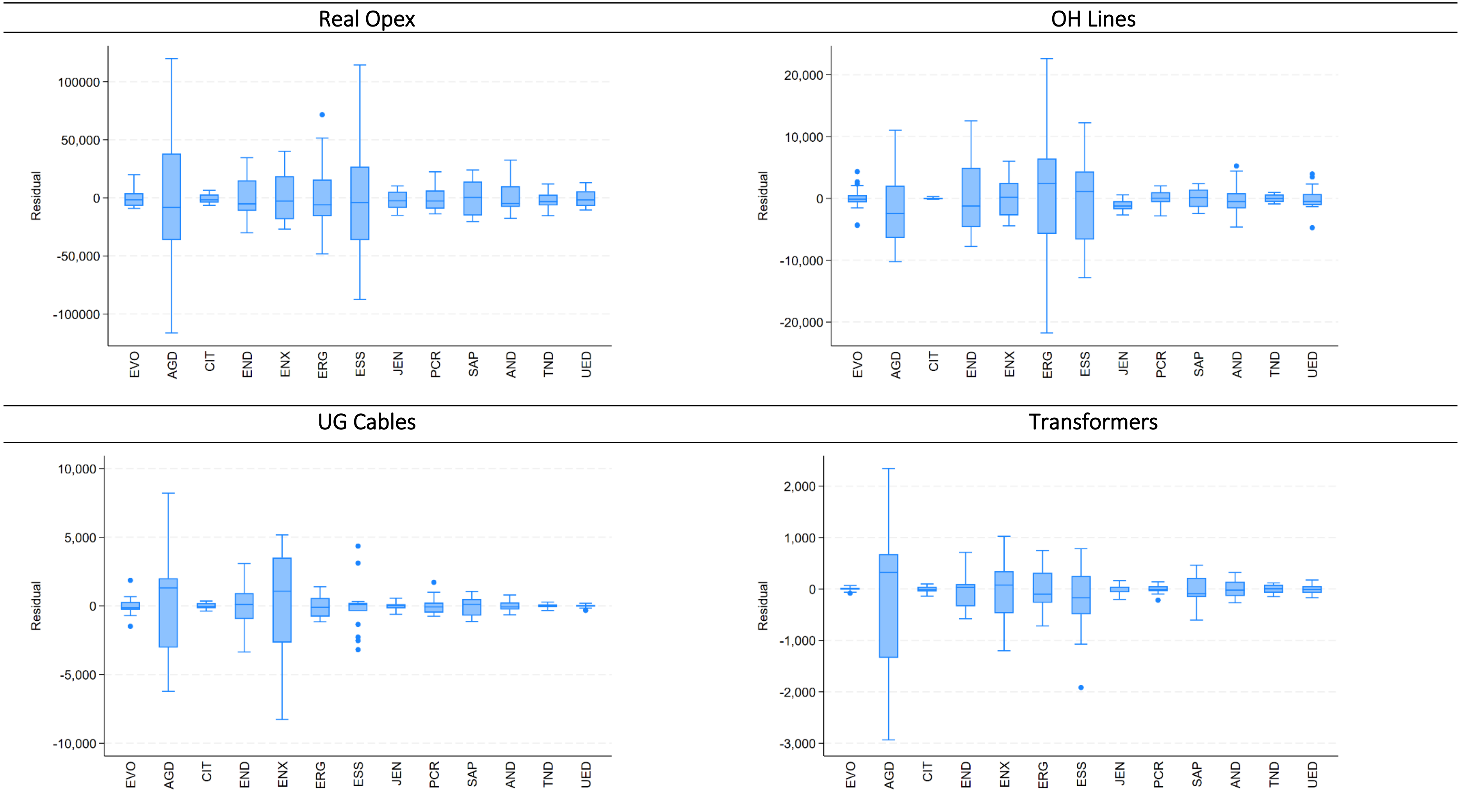


Table 4.1 R^2 Values Comparisons

	*EI 2006-2018**				*2006-2023***				*2006-2024*			
	Opex	*OH Lines*	*UG Cables*	*Transfor.*	*Opex*	*OH Lines*	*UG Cables*	*Transfor.*	*Opex*	*OH Lines*	*UG Cables*	*Transfor.*
EVO	0.143	0.925	0.975	0.990	0.018	0.853	0.962	0.985	-0.003	0.886	0.997	0.988
AGD	0.256	0.439	0.891	0.928	0.460	0.316	0.827	0.800	0.508	0.266	0.806	0.726
CIT	0.539	0.461	0.982	0.980	0.530	0.068	1.000	0.984	0.528	0.034	0.965	0.986
END	0.313	0.090	0.984	0.978	0.141	0.429	0.984	0.971	0.182	0.502	0.982	0.972
ENX	0.674	0.958	0.989	0.995	0.539	0.886	0.942	0.975	0.613	0.966	0.904	0.961
ERG	0.132	0.460	0.980	0.976	0.461	-4.778	0.947	0.957	0.189	0.373	0.949	0.959
ESS	0.088	0.962	0.865	0.985	0.127	0.940	0.800	0.853	0.083	0.940	0.825	0.850
JEN	0.738	0.711	0.989	0.976	0.162	0.845	0.993	0.980	0.079	0.572	0.992	0.973
PCR	0.408	0.682	0.949	0.997	0.488	0.852	0.975	0.997	0.561	0.873	0.981	0.997
SAP	0.816	0.713	0.990	0.975	0.689	0.693	0.967	0.962	0.731	0.727	0.972	0.966
AND	0.797	0.877	0.996	0.986	0.540	-24.129	0.997	0.971	0.582	0.677	0.996	0.974
TND	0.249	0.944	0.976	0.974	0.231	0.970	0.977	0.969	0.150	0.974	0.980	0.971
UED	0.332	0.871	0.993	0.986	0.232	0.847	0.994	0.989	0.212	0.851	0.995	0.990
AVG	0.422	0.699	0.966	0.979	0.355	0.700***	0.951	0.953	0.340	0.665	0.949	0.947

Note: *From Economic Insights (2020b: Appendix A). ** Using the ABR25 dataset, restricted to 2006-2023. *** Excluding the two negative R^2 values.

4.4 Effect of Alternative Starting Values

This section examines the input demand model goodness-of-fit and residuals when the coefficients from the maximum likelihood (ML) models are used as starting values. Table 4.2 reports the R^2 values for each DNSP and input across the different datasets, now re-estimated using the ML coefficients as starting values.

Table 4.2 R^2 Values using different datasets (ML as starting values)

	2006-2023				*2006-2024*			
	Opex	*OH Lines*	*UG Cables*	*Transfor.*	*Opex*	*OH Lines*	*UG Cables*	*Transfor.*
EVO	0.039	0.853	0.962	0.985	0.036	0.886	0.905	0.988
AGD	0.462	0.316	0.830	0.800	0.508	0.312	0.806	0.726
CIT	0.530	0.068	0.970	0.984	0.547	0.034	0.965	0.986
END	0.147	0.429	0.984	0.971	0.182	0.502	0.982	0.972
ENX	0.539	0.965	0.942	0.975	0.613	0.966	0.904	0.962
ERG	0.461	0.354	0.947	0.957	0.189	0.373	0.949	0.959
ESS	0.127	0.940	0.800	0.853	0.083	0.940	0.825	0.864
JEN	0.162	0.845	0.993	0.980	0.118	0.843	0.992	0.973
PCR	0.488	0.852	0.977	0.997	0.563	0.873	0.981	0.997
SAP	0.703	0.693	0.967	0.963	0.735	0.727	0.972	0.966
AND	0.540	0.756	0.997	0.971	0.582	0.677	0.996	0.974
TND	0.231	0.970	0.977	0.970	0.241	0.974	0.980	0.973
UED	0.233	0.847	0.994	0.989	0.212	0.851	0.995	0.990
AVG	0.359	0.684	0.949	0.954	0.355	0.689	0.942	0.948

Note: Results were obtained using the ABR25 dataset, restricted to the relevant years as shown.

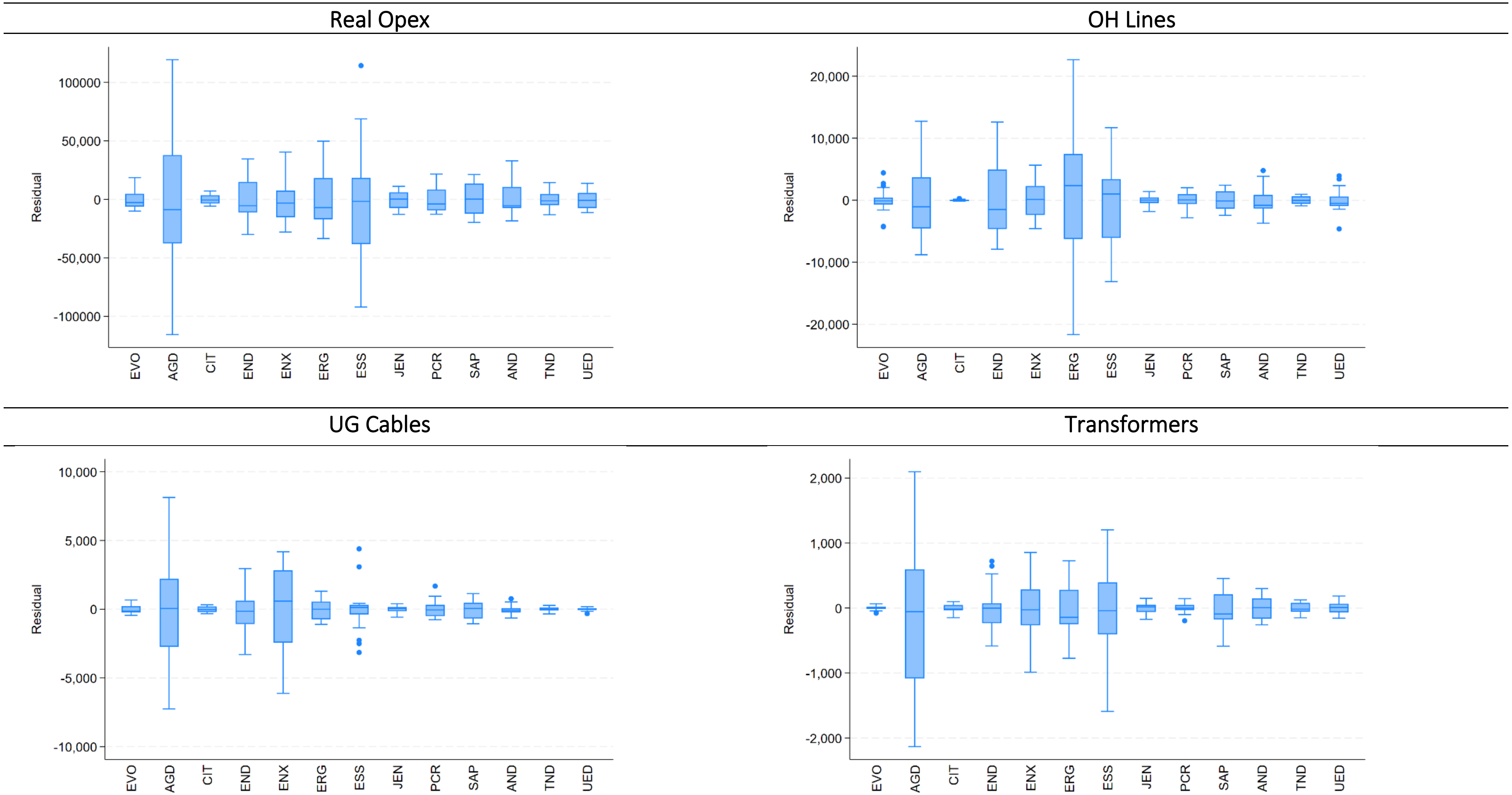
Figure 4.2 presents the boxplots of residuals associated with these models. It shows that using the ML estimates as starting values resolves the issue of residuals not being centred around zero for the ERG and AND OH lines regressions.

The main observations on the models that use ML starting values are:

- *2006–2023 data sample*: On average, the R^2 values are closely aligned with those obtained using 0.001 as the starting values. Importantly, there are no negative R^2 values. Again, the issue with the OH lines input for ERG and AND has disappeared.
- *2006–2024 data sample*: The R^2 values are generally consistent with those from the 2006–2024 estimations using 0.001 as the starting value. As with that specification, no unusual R^2 values are observed.

These results lend support to the use of the ML estimates of coefficients as the starting values in next estimating the nonlinear input demand models.

Figure 4.3: Residual Distributions by DNSP (2006-2023, ML as starting values)



Note: Results were obtained using the ABR25 dataset, restricted to 2006-2023.

4.5 Conclusions

This section highlights that in the Leontief models restricted to 2006-2023 period, on which the output weights used in the benchmarking report were based, two of the OH lines models have problematic residuals that are not centred near zero. The same models have a negative R^2. This problem is resolved by using the 2006-2024 period. Alternatively, it is resolved by using ML-based starting values instead of the standard starting values.

The negative R^2 appears to have occurred because the model's predicted values performed worse than the simple average of the observed values. This issue seems to have arisen from computational instability, as it was resolved when alternative starting values were used or when one additional year of data was included in the estimation.

These results, and the results presented in section 2, tend to suggest that including the 2024 data improves the model's performance compared with using the sample restricted to 2006-2023 for estimating the output weights.

5 Other Issues Raised by Stakeholders

Sections 3 and 4 considered some issues raised by Frontier Economics. This section discusses some other comments made by stakeholders.

5.1 Stakeholder Comments and our response

AusNet argued that the substantial changes in output weights estimated using the 2006-2023 sample, compared to those estimated with the 2006-2018 sample, "raise concerns about the stability and reliability of the underlying econometric modelling" and "the model specification and estimation methods make the estimated output weights highly sensitive to changes in input data". SAPN expressed similar views. Section 2 examines the sources of the changes in output weights. Importantly, the data sample has increased from 13 to 18 observations per regression, a roughly 40 per cent increase. And there has also been a major change in data definitions relating to CCOs. Although it is desirable to have stability in the output weights, it may be unrealistic to expect only minor change in these circumstances.

SAPN points to the fact that in many or most of the input demand functions, of which there are four for each DNSP and 52 in total, some of the coefficients pertaining to outputs are zero or close to zero. Sometimes, these models only identify a single cost driver for a particular input and DNSP. SAPN draws the conclusion that "none of the estimated Leontief models are reliable. Therefore, the basis for the updated output weights proposed by Quantonomics should be rejected".

We do not agree with this reasoning. There is no problem with coefficients taking zero values in the individual input demand functions. As discussed in section 3.3 in response to Frontier Economics, this is a mistaken argument. The purpose and advantages of estimating the models separately for individual DNSPs were explained by CEPA (2024: 14), namely that it allows a highly flexible estimation procedure. The objective being to use the estimated parameters to derive the proportionate importance of each output as a driver of each input cost, and ultimately total cost, at the *industry* level. The reasonableness of this set of models needs to be evaluated at the industry level, where it is applied to construct weights. As shown in Tables 2.8 to 2.11, at the aggregate level, each input has a different mix of cost drivers. None of the outputs has zero effect on costs for any input. SAPN has not indicated how much of the cost of each input it expects should be determined by each output, so it is not possible to conclude whether those results are consistent or inconsistent with its expectations.

Various modelling suggestions are made by submitters:

(a) EVO notes that the input demand models "do not account for time-varying inefficiency". SAPN concurs that they "do not allow for inefficiency as a possible explanator of DNSPs' inputs … beyond a smooth linear trend". SAPN therefore recognises that the time-trend component of the model could capture time-varying inefficiency. Given the presence of the time trend, which will capture the average rate of all systematic time-varying effects, EVO and SAPN have not made clear how the models do not allow for time-varying inefficiency.

(b) EVO states that the models omit "some potentially relevant output variables (such as for consumer energy resources)". SAPN also criticises the output specification for failing to include "the delivery of CER services". AGD notes that "demand-sides initiatives and CER integration … may appear as output reductions …". Whilst it is desirable to account for these increasingly important services, the AER investigated this issue in 2023 and concluded that at the present time there is insufficient data include export services in the benchmarking framework, but it "may reconsider this position in the future when more robust export services expenditure data is available" (AER, 2023: 8).

(c) EVO argues that the models should be estimated using panel regressions, rather than by individual DNSP. "As the models are not estimated as panel regressions, all data variation is derived exclusively from variation over time (i.e., there is no cross-sectional variation), which results in poorer model fit." While this observation is not unreasonable, such an approach would reduce the flexibility of the modelling, noting that CEPA regarded this as an advantage of the Economic Insights method.

(d) AND proposes that smoothing techniques be used to minimise abrupt changes in output weights due to model sensitivity. We have addressed the reasons for changes in output weights. Note that for the last four samples from 2006-2021 to 2006-2024 there have not been major changes in the output weights. Regarding smoothing techniques, while nothing specific was proposed, there is a wide range of methods, and such an approach may introduce *ad hoc* use of such methods, and subjective judgements. If that were the case, it would not be suitable.

(e) SAPN suggests that an alternative method of averaging the results across DNSPs be considered. This question is explored in section 5.2.

Some have expressed views supportive of the output weights analysis. JEN stated:

> We recognise the trade-offs between model complexity and statistical significance. Given the current data limitations, we agree with Quantonomics' view to retain the existing estimation approach. It offers regulatory certainty and consistency, and the results are broadly aligned with those suggested by CEPA's independent review.

Overall, there may be merit in further investigation of methods of estimating output weights, although we note that CEPA investigated the output weights methodologies in 2024 in its independent review and associated industry consultation. Hence, there is no presumption that a consistently better method could readily be found.

Differing views were expressed on the frequency of updating output weights. JEN expressed support in "revisiting the output weight estimation approach in five years, when a larger data sample becomes available". ENX/ERG "recommend that the AER adopt an ongoing (annual) update process for non-reliability output weights". There are pros and cons to each suggestion. One benefit of an annual updating process is that it would reduce the likelihood of substantial data revisions accumulating over time and causing a large impact when the output weights are updated. On the other hand, updating every year may be less amenable to making year-to-year comparisons of changes in productivity and rankings on a like-for-like basis.

5.2 Alternative Methods of Averaging of Output Cost Shares

According to SA Power Networks:

> A further limitation of the current method is that the industry non-reliability output weights are calculated based on the 'total cost shares' across the 13 DNSPs, which causes the results to be dominated by the largest DNSPs. This happens because the largest DNSPs contribute the most to the total estimated industry cost. There is little justification for assuming that the relationship between inputs and outputs for the largest DNSPs should dictate the output weights applied to all other DNSPs. By contrast, the econometric benchmarking models used by the AER to assess the efficiency of base year opex effectively weight all DNSPs equally, which in our view is more appropriate. (SA Power Networks, 2025: 8)

It is correct that in Economic Insights' method, the industry output cost shares are a weighted average of the output cost shares of the DNSPs, where the weights are their shares to total industry cost. This method has a clear internal logic. The method is designed to fully allocate each DNSP's cost of using each input to each of the four outputs. The cost attributed to one output can then be aggregated over inputs and over DNSPs to obtain the share of that output in total industry cost.

Whilst noting the logic of the Economic Insights method, given the substantial changes in the estimated output weights compared to those previously estimated using the 2006-2018 data sample, it is of interest to consider whether an alternative averaging method yields different weights or patterns of change. The alternative examined in this section is to firstly calculate the output cost shares for each input for each DNSP using Economic Insights' method, and then to average the cost shares of each output over DNSPs using a simple (arithmetic) average of the cost shares of the DNSPs instead of a weighted average.

The results of this approach (using standard starting values for parameters in the NLS models) are shown in Table 5.1 and Figure 5.1. The corresponding results under weighted averaging are presented in Table 2.2 and Figure 2.1 in section 2. The arithmetic averaging of the output cost shares results for the 2006-2023 data sample show that Energy Throughput has a much higher estimated cost share, namely 15.9 per cent compared to 10.8 per cent using the weighted average cost shares. RMD has a much lower cost share of 38.3 per cent, compared to 47.8 per cent using the weighted average cost shares. Customer numbers has a cost share of 15.8 per cent, similar to the 15.2 per cent weighted average cost share. Circuit length has a share of 30.0 per cent, which is higher than its 26.2 per cent weighted average cost share.

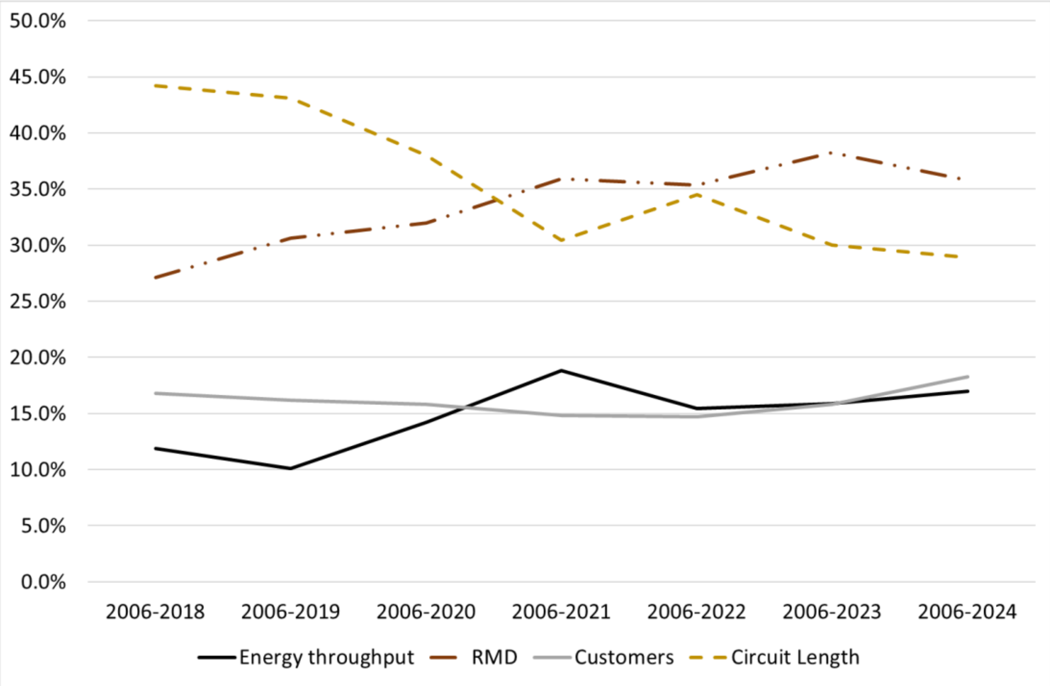
Figure 5.1 shows a similarity of trends over different data samples. The key changes are: (a) a substantial increase in the cost share of RMD between the 2006-2018 and 2006-2023 data samples; (b) a substantial decrease in the cost share of Circuit Length between the same two data samples; and (c) Energy Throughput also has a significant increase in its average cost share. Overall, the nature of the changes on the cost shares between the 2006-2018 and 2006-2023 data samples are broadly directionally similar to the weighted average method.

Table 5.1 DNSP Simple Average Output Cost-Share Weights with Different Data Samples

	2006-2018	*2006-2019*	*2006-2020*	*2006-2021*	*2006-2022*	*2006-2023*	*2006-2024*
Energy Throughput	11.9%	10.1%	14.2%	18.8%	15.5%	15.9%	17.0%
RMD	27.1%	30.6%	32.0%	35.9%	35.3%	38.3%	35.8%
Customers	16.8%	16.2%	15.8%	14.8%	14.7%	15.8%	18.3%
Circuit Length	44.2%	43.1%	38.0%	30.5%	34.5%	30.0%	28.9%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: Results were obtained using the ABR25 dataset, restricted as shown.

Figure 5.1 DNSP Simple Output Cost-Share Weights with Different Data Samples



6 Conclusions

This memo has examined the drivers of changes in the output weights and addressed the criticisms raised by DNSPs regarding model specification.

As discussed in Section 2, most of the changes in the output weights were driven by the inclusion of additional data, with revisions to data definitions also playing an important role. Specifically, the sharp increase in the cost share of RMD reflects the combined effect of data revisions and the addition of new data, both of which substantially increase its weight. The modest increase in the cost share of Energy Throughput is due to largely offsetting effects between data revisions and new data. The sharp decline in the cost share of Circuit Length is mainly explained by the additional data, which reduced the estimated influence of this output on input demands.

The largest changes in output weights, particularly for RMD and Circuit Length, occurred with the inclusion of the 2020 and 2021 data. With the additional data, the relationship between Opex and Circuit Length weakened and the importance of Circuit Length as a driver of OH Lines and UG Cables declined (perhaps reflecting changes (increases) in line capacities). Costs were increasingly attributed to RMD across all inputs. At the same time, the sharp fall in capital input prices in 2020 and 2021 raised the share of Opex in total costs. Since RMD accounts for a relatively high share of Opex costs while Circuit Length accounts for a relatively low share, these effects accentuated the overall shift in cost shares away from Circuit Length and towards RMD.

In Section 3, we consider Frontier Economics' proposal to use alternative starting values, obtained by adding an extra step to the estimation process. We recognise the merit of their approach and recommend it as a potential improvement. However, we also highlight its limitations, including convergence failures in two of the 52 models. For these cases, we suggest reverting to the standard starting value of 0.001. While the approach of using ML-derived starting values produces similar weights for Circuit Length and Customer Numbers and lower weights for Energy Throughput, it further amplifies the increase in the RMD weight.

In Section 4, we review the residuals and model fit of the tested models. For ERG and AND in the overhead lines estimation, residuals were not centred, and the R^2 values were highly negative, indicating computational issues. These problems were resolved by either including the 2024 data or by adopting the ML-based starting values suggested by Frontier Economics. As the inclusion of the 2024 data does not materially affect the output weights, we conclude that these issues do not undermine the weight estimates. Apart from these exceptions, the R^2 values are broadly consistent across all datasets and under the ML approach, with lower R^2 values for opex input demand functions, moderate values for overhead lines, and very high values for underground cables and transformers.

In Section 5, we consider the remaining stakeholder comments. An important concern of several DNSPs is the size of the changes in the updated output weights, compared to those estimated by Economic Insights in 2020. While stability in output weights is desirable, given the substantial additional data and data revisions since that previous study, it may be unrealistic to expect only minor changes. The results presented in section 2 for successive results using the data samples from 2006-2021 to 2006-2024 do not support the contention of some stakeholders that the Leontief method suffers from instability of estimated output weights. In response to some of the detailed criticisms of specific input demand models, we observe that the models were designed to provide flexibility and to estimate weights at the *industry* level rather than for individual DNSPs.

We have addressed SAPN's claim that the output weights are disproportionately influenced by the largest DNSPs by testing an alternative averaging method. We do not agree with the logic of SAPN's approach, given the interest in industry cost shares. However, for completeness, we tested SAPN's alternative method. This resulted in a much higher estimated cost share for Energy Throughput, a lower share for RMD, a similar share for Customer Numbers, and a higher share for Circuit Length. Nevertheless, the overall trends in output weight changes were broadly directionally similar to those using the Economic Insights method.

In summary, while there may be scope to refine the models used to estimate output weights, the evidence shows that the material changes in weights are primarily driven by data inclusion and revisions. The observed changes in output weights are not an artefact of estimation shortcomings, and potential improvements in model specification do not alter the materiality of these results.

Appendix A Actual Values *versus* Fitted Values, 2006-2023 Data

Figures A.1 to A.13 show the actual values (red dots) and the fitted or predicted values (black line) from the output weight estimations using the 2006–2023 period and ABR25 dataset applied in the Benchmarking Report 2025. The key observations are:

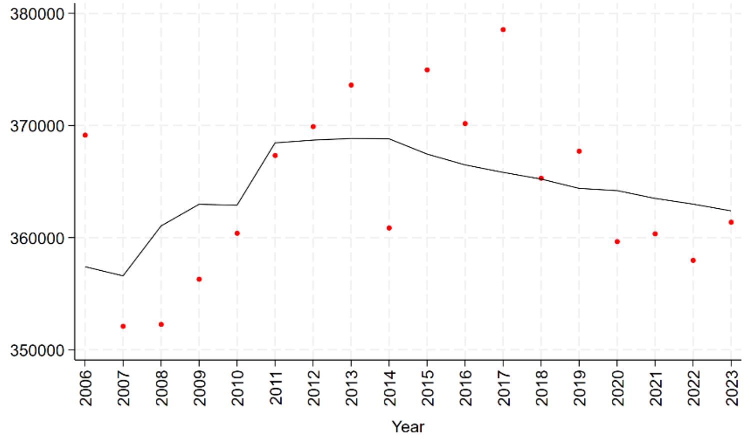
- Underground cables and transformers: The fitted values generally align closely with the actual values. This is consistent with the boxplot analysis, where the medians of the residuals for these variables are near zero.
- Opex: The alignment between actual and predicted values is weak overall, and this may reflect a model misspecification. While addressing this is beyond the scope of the current analysis, it is worth noting that the model only incorporates a linear trend, even though the opex trend commonly shifts mid-sample due to efficiency improvements. Allowing for a quadratic trend could be a useful refinement.
- Overhead lines: The alignment is strong for some DNSPs (EVO, ENX, ESS and TND) but noticeably weak for others (AGD, CIT and END). More importantly, for two DNSPs (ERG and AND), the fitted line does not intersect the cloud of actual values, suggesting computational issues.

Figure A.1: EVO – Actual *vs* Predicted Values (2006-2023)

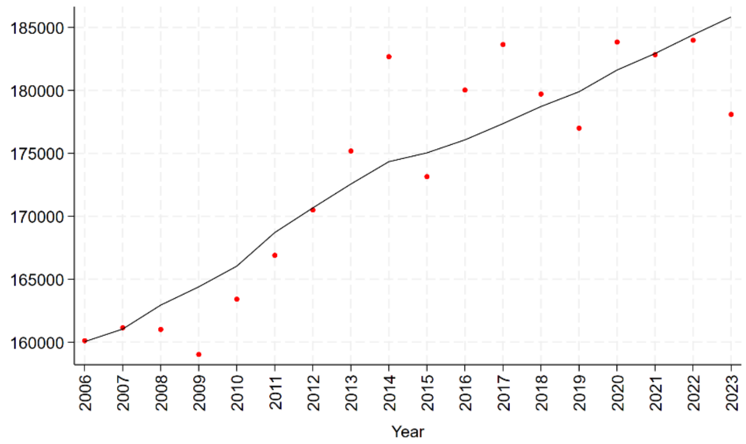
Figure A.2: AGD– Actual vs Predicted Values (2006-2023)

Real Opex

OH Lines



UG Cables



Transformers

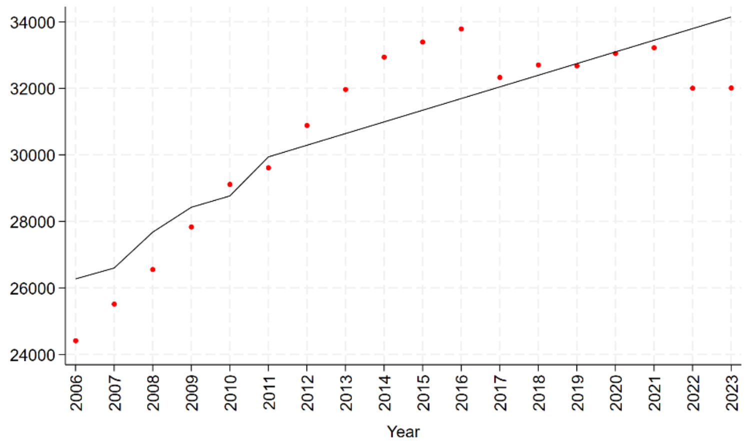


Figure A1.3: CIT– Actual vs Predicted Values (2006-2023)

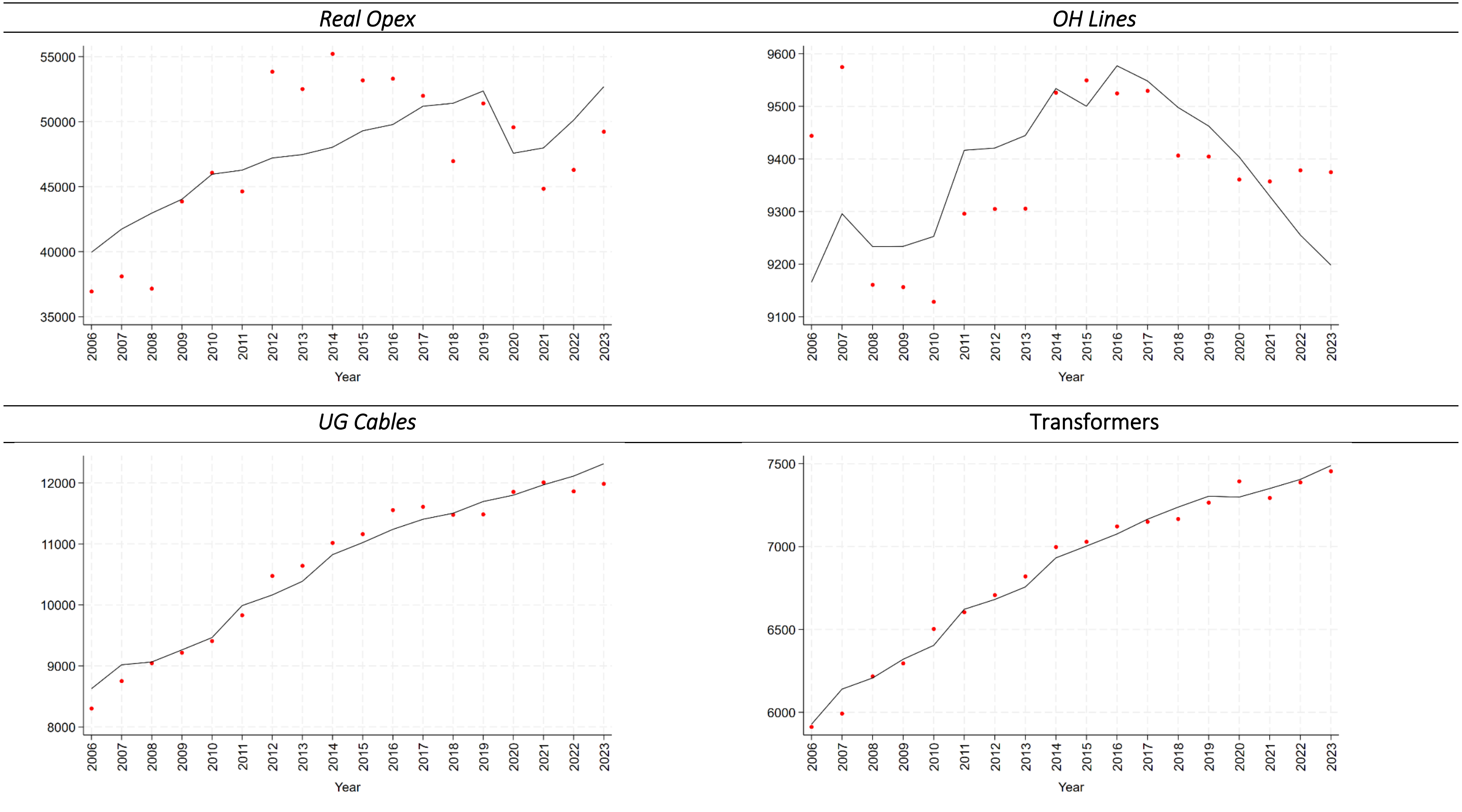


Figure A1.4: END– Actual vs Predicted Values (2006-2023)

Quantonomics

QUANTITATIVE ECONOMICS

Figure A1.5: ENX– Actual vs Predicted Values (2006-2023)

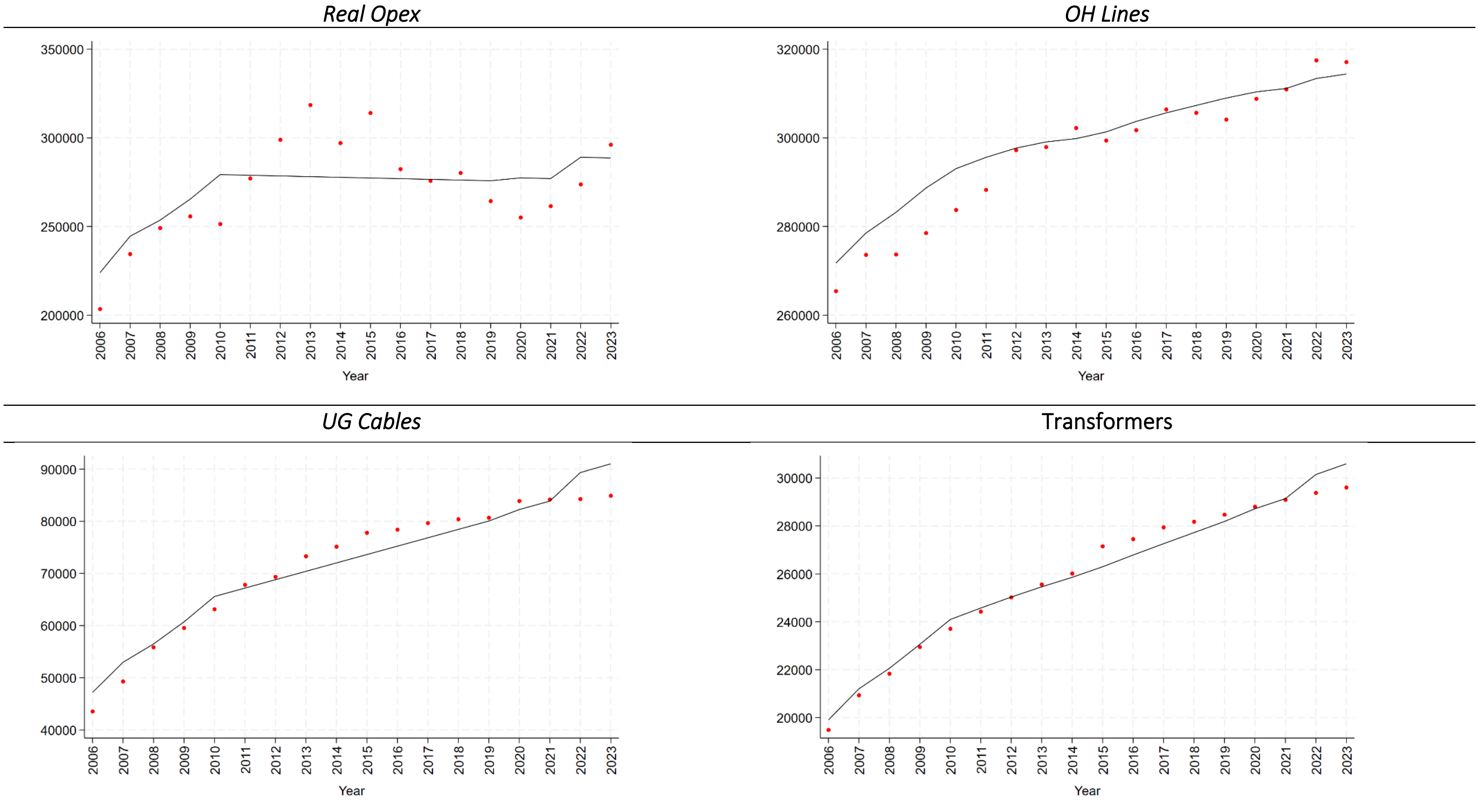
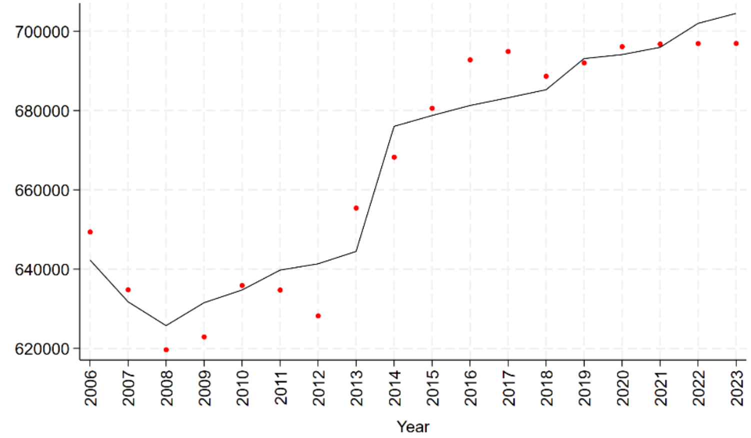


Figure A1.6: ERG– Actual vs Predicted Values (2006-2023)

Figure A1.7: ESS– Actual vs Predicted Values (2006-2023)

Real Opex

OH Lines



UG Cables

Transformers

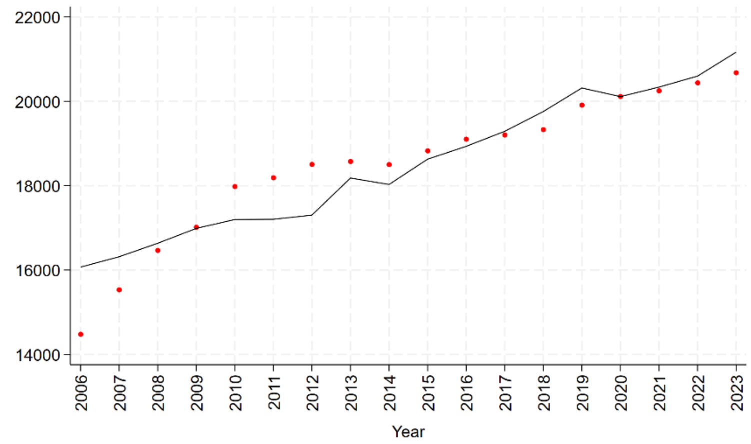
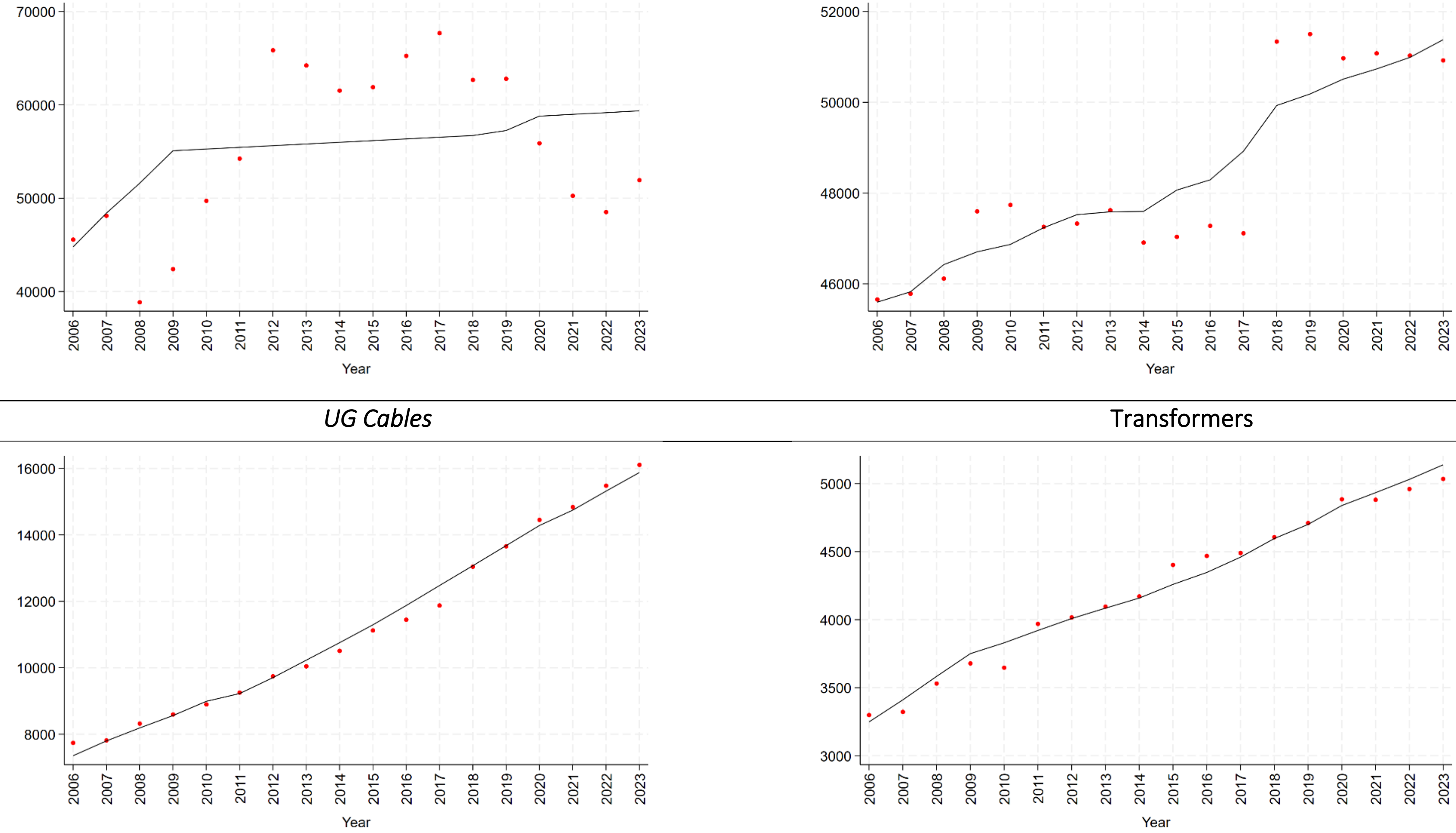


Figure A1.8: JEN– Actual vs Predicted Values (2006-2023)

Real Opex | *OH Lines*



UG Cables | Transformers

Figure A1.9: PCR– Actual vs Predicted Values (2006-2023)

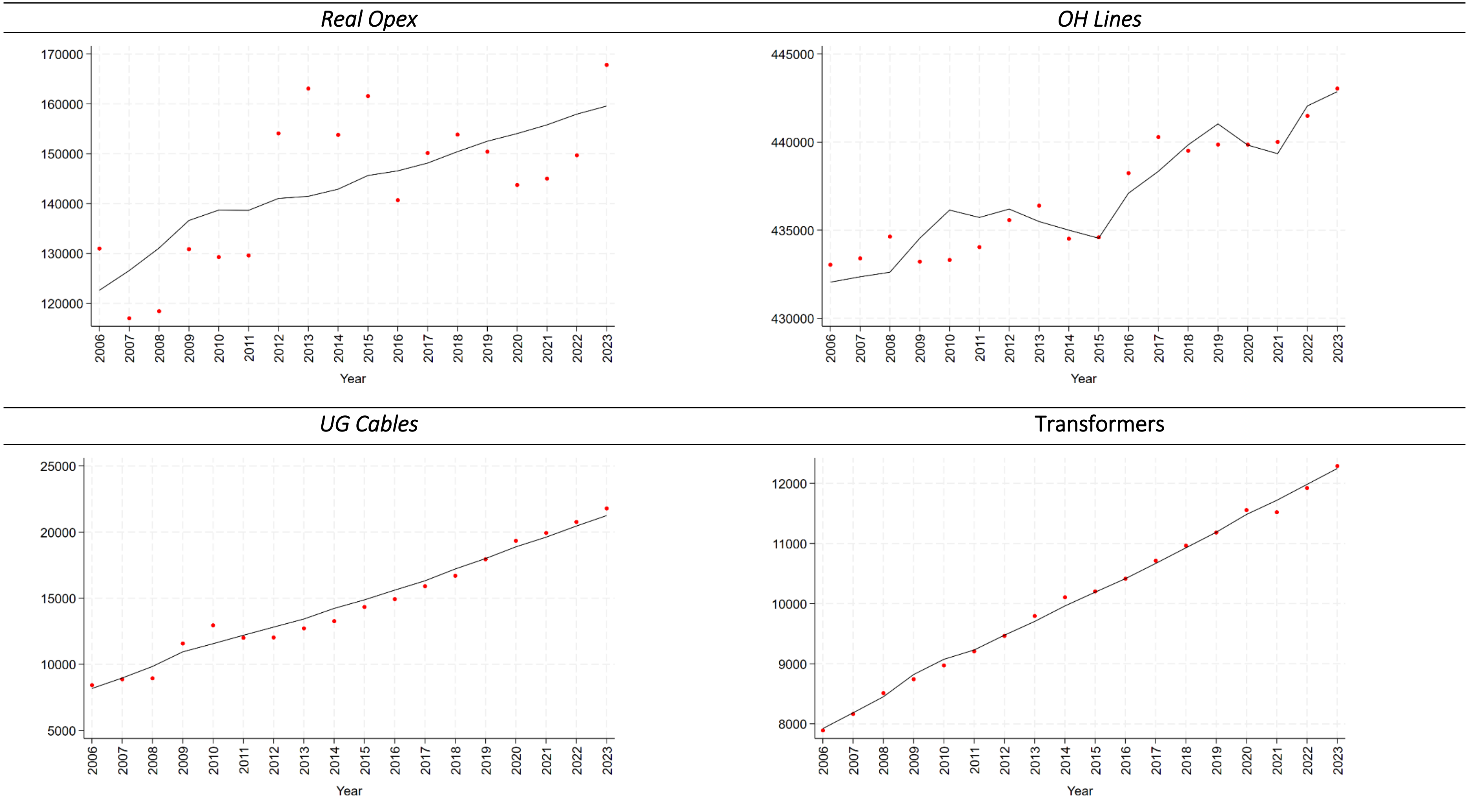
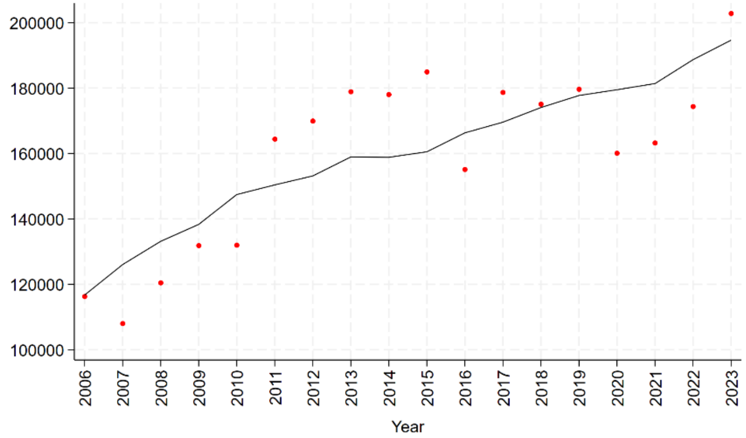
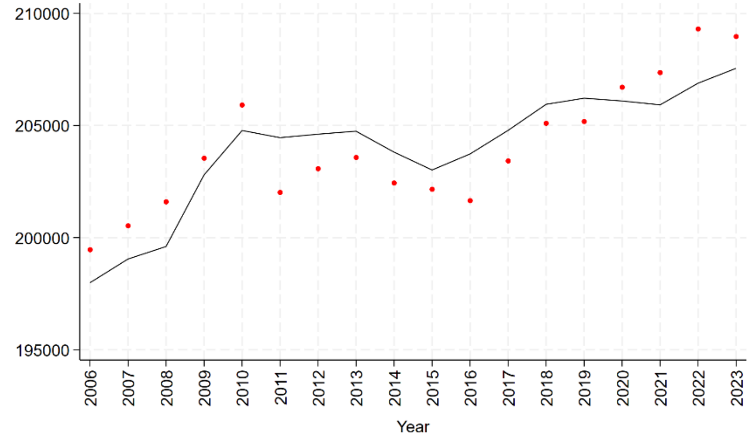


Figure A1.10: SAP– Actual vs Predicted Values (2006-2023)

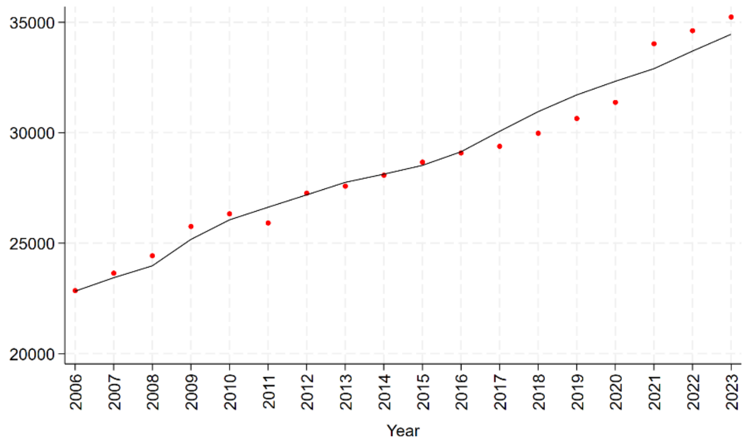
Real Opex



OH Lines



UG Cables



Transformers

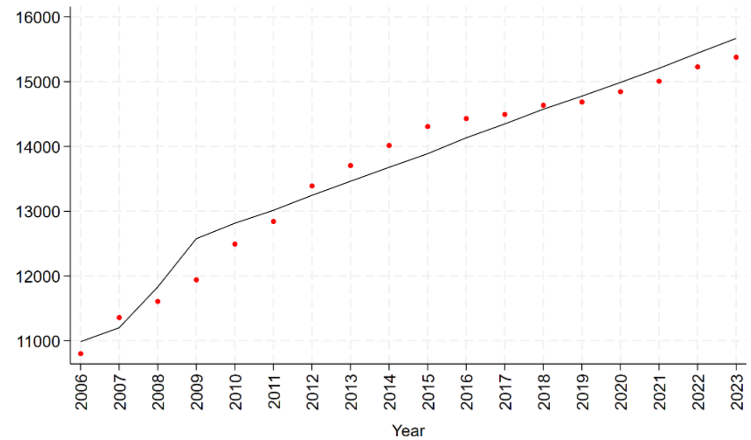
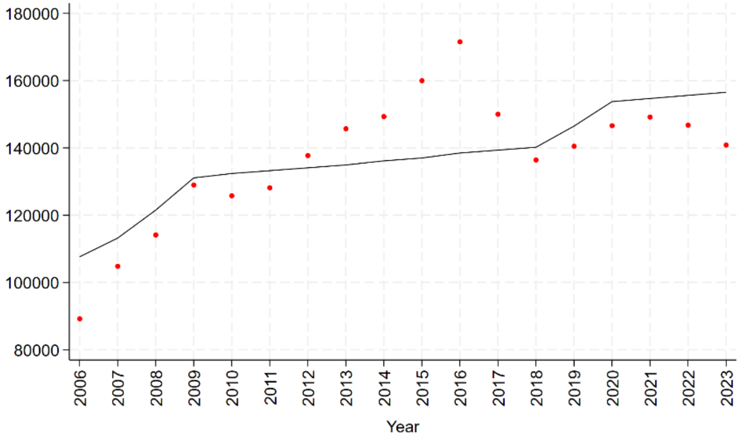
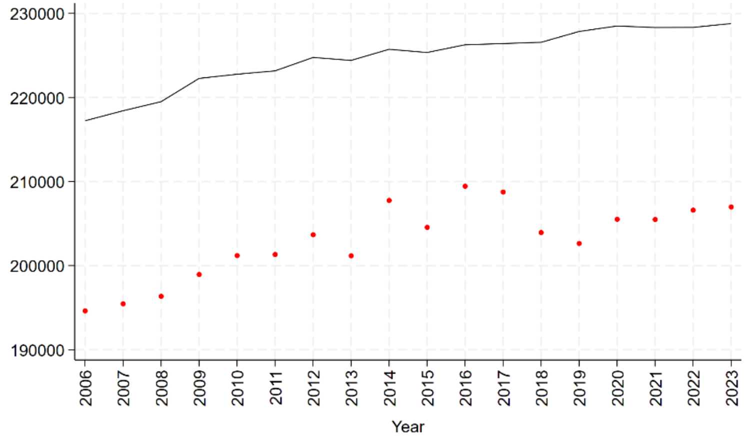


Figure A1.11: AND– Actual vs Predicted Values (2006-2023)

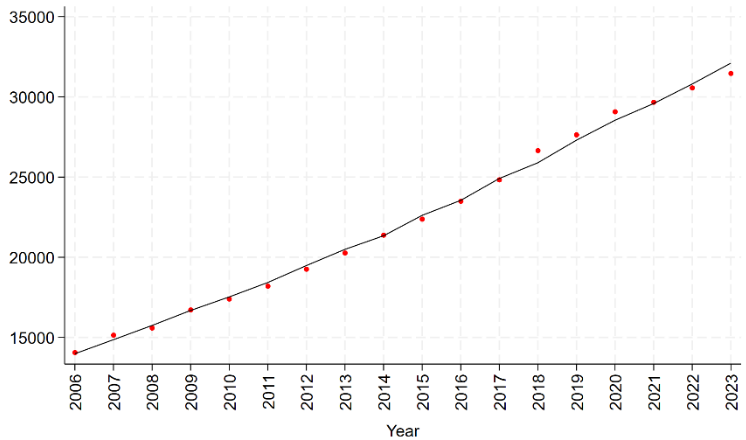
Real Opex



OH Lines



UG Cables



Transformers

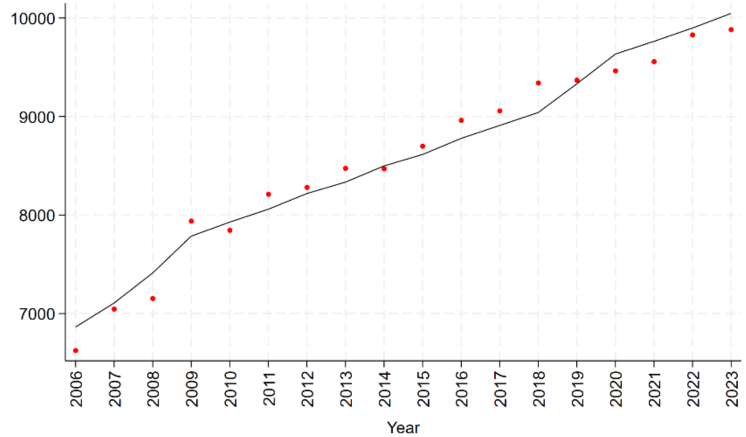
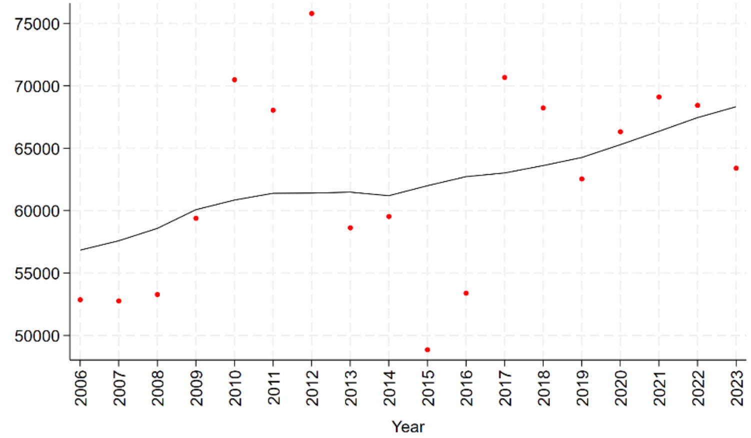
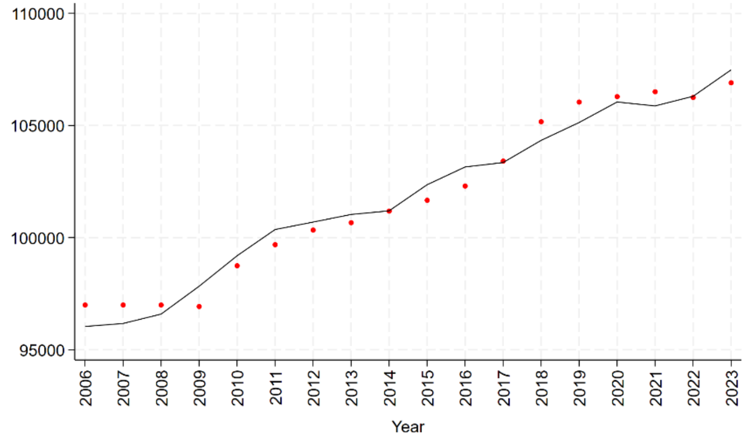


Figure A1.12: TND– Actual vs Predicted Values (2006-2023)

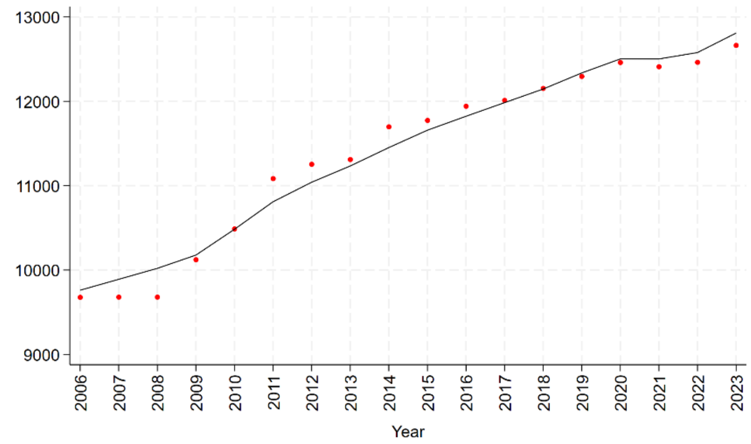
Real Opex



OH Lines



UG Cables



Transformers

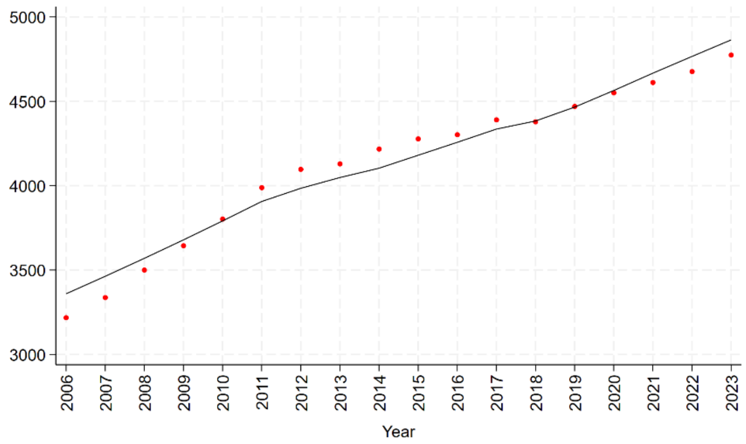
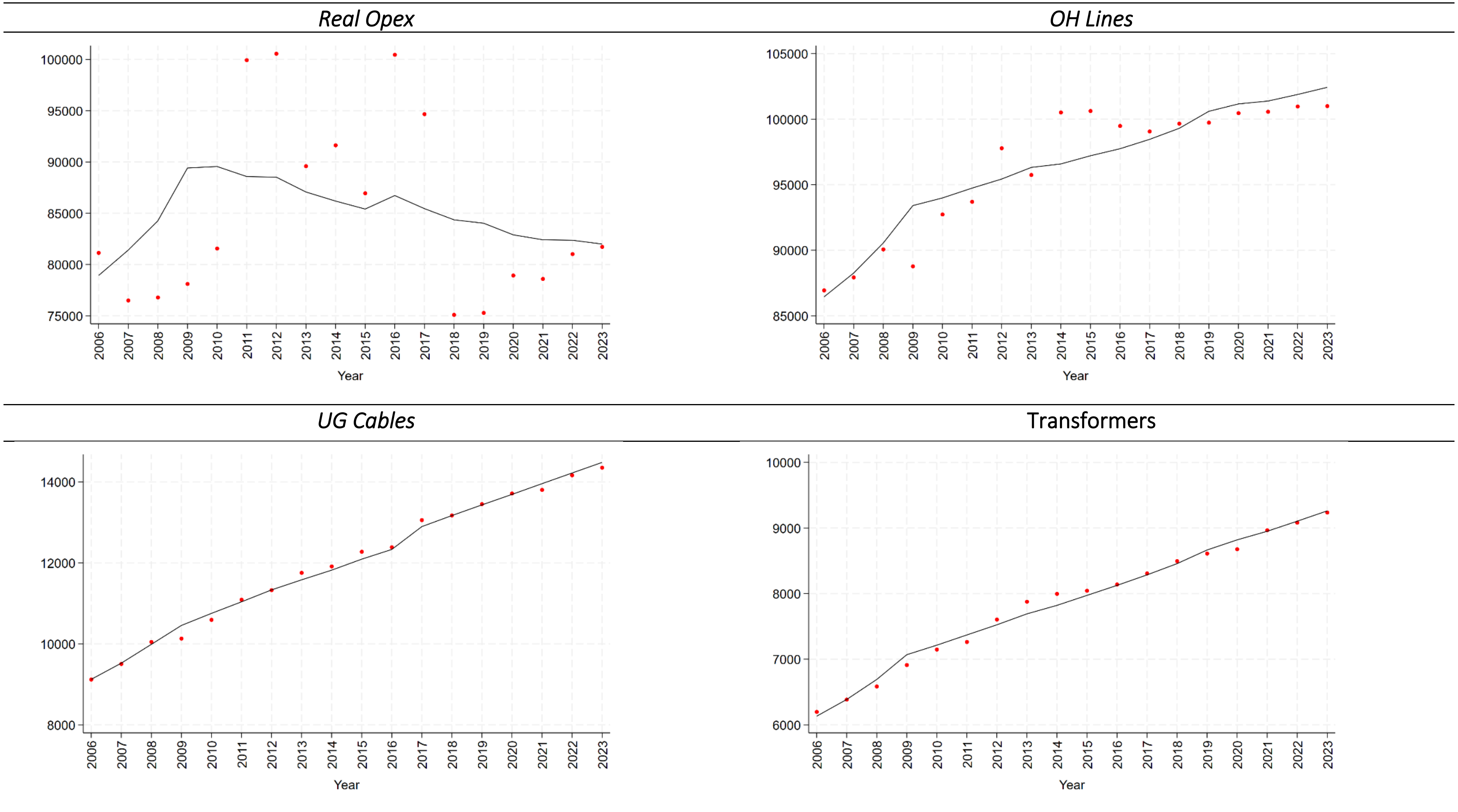


Figure A1.13: UED– Actual vs Predicted Values (2006-2023)



References

Australian Energy Regulator (AER) (2023) *Incentivising and measuring export service performance*. Final Report, March.

Baum CF (2006) *An Introduction to Modern Econometrics Using Stata*. Stata Press.

Coelli T, Rao P, O'Donnell C, et al. (2005) *An Introduction to Efficiency and Productivity Analysis*. 2nd edn Springer.

Economic Insights (2020a) *Economic Benchmarking Results for the Australian Energy Regulator's 2020 DNSP Annual Benchmarking Report*. Prepared for Australian Energy Regulator by Denis Lawrence, Tim Coelli and John Kain, October.

Economic Insights (2020b) *Review of reports submitted by CitiPower, Powercor and United Energy on opex input price and output weights*. 18 May.

Evoenergy (2025) Submission: 2025 DNSP Annual Benchmarking Report – Draft report and results.

Peyrache A (2024) *Review of the AER's estimated non-reliability output weights used in the TFP and MTFP benchmarking models*. Centre for Efficiency and Productivity Analysis (CEPA).

SA Power Networks (2025) Submission: Feedback on Quantonomics' update of non-reliability output weights.