Response to Phase 2 consultation on econometric benchmarking models

A report prepared for Evoenergy | 27 March 2026

Contents

Tables

Figures

1 Executive summary

1.1 Background

On 26 November 2025, the Australian Energy Regulator (AER) published a report by its adviser on expenditure benchmarking, Quantonomics, as part of Phase 2 of the AER's consultation on the specification of the econometric models used to benchmark the operating expenditure (opex) of Distribution Network Service Providers (DNSPs).1

The Phase 2 report builds on an earlier report prepared by Quantonomics as part of Phase 1 of the models that the AER currently uses to benchmark DNSP opex are mis-specified, and explored different treatments of jurisdictional time trends as a way of improving the specification and statistical performance of the models.

In response to Quantonomics' Phase 1 report, we prepared a report that argued that the source of the mis-specification problem was likely to be that the existing benchmarking models assume DNSPs' levels of efficiency remain invariant over time.[2] But, in fact, there is compelling evidence that many DNSPs have become more efficient over time since the AER began using benchmarking analysis in 2015 to assess the efficiency of opex forecasts, when setting revenue allowances. Since the existing models are incapable of accounting for this, the resulting opex cost functions will be mis-estimated, and the AER's assessment of efficient opex will be unreliable.

We concluded that allowing for different time trends for different jurisdictions alone would not address the source of the mis-specification problem. We therefore supported Quantonomics' recommendation to the AER that time-varying inefficiency models should be explored, as a possible way to address the model mis-specification problem, and we made some suggestions (i.e., classes of time-varying inefficiency models and statistical software packages) that could be used to do this.

The focus of Quantonomics' Phase 2 report is to investigate several different classes of time-varying inefficiency models, of which three receive the most attention:

- the Battese and Coelli (1995) (BC95) model;
- the Kumbhakar (1990) (Kumb90) model; and
- the least squares econometrics (LSE) model.

Quantonomics tested several variants of each of these models, including by:

- considering different treatments of jurisdictional time trends and time trends for individual DNSPs;
- the inclusion of a general technical change (GTC); and

[1] Quantonomics, Electricity distribution benchmarking opex model development, 26 November 2025. (Quantonomics Phase 2 report).

[2] Frontier Economics, Response to AER's Phase 1 consultation on potential mis-specification issues – electricity distribution opex cost function, 27 February 2025.

- the use of a 'half normal' (rather than the more general truncated normal) assumption for the distribution of efficiency for the stochastic frontier models.

Quantonomics fitted these models to the historical Regulatory Information Notice (RIN) benchmarking data between 2006 and 2023 and presented results for each model.

Quantonomics then assessed the performance of the models against various selection criteria and identified four models that it considered perform best according to those criteria:

- one BC95 model;
- one Kumb90 model; and
- two LSE models.

Quantonomics stated that:

> *Overall, these models produce results that are consistent with expected output elasticities and the correct signs of coefficients, particularly in the longer-period sample. They also exhibit few or no monotonicity violations, a high level of goodness of fit, satisfactory performance in statistical tests, and stable efficiency scores.*[3]

However, Quantonomics stopped short of recommending that these models should replace the AER's existing econometric models, even though (according to the results presented by Quantonomics) they objectively outperform the existing models. Quantonomics stated that:

> *...while these findings highlight the potential value of incorporating time-varying inefficiency models into the opex assessment framework, they remain preliminary. We reiterate that these are initial views and further analysis is required.*[4]

We have been asked by Evoenergy to:

- investigate and test the analysis presented in Quantonomics' Phase 2 report;
- advise which time varying inefficiency models, if any, should replace the AER's existing benchmarking models; and
- advise how the results from those models should be used to estimate an efficient level of base-year opex.

3 Quantonomics Phase 2 report, p. 110.

4 Quantonomics Phase 2 report, p. 143.

1.2 Key findings and recommendations

Several issues with Quantonomics' analysis needed to be addressed before the models could be evaluated

We identified a number of issues with the Quantonomics' analysis that needed to be addressed before the time-varying inefficiency models could be assessed properly:

- Quantonomics' treatment of the efficiency time trend in some BC95 models that results in estimates of expected efficiency that are essentially time invariant;
- The restrictive nature of the efficiency time trend in the Kumb90 models that tend to produce highly implausible estimates of efficiency over time;
- The restrictive assumption in LSE models that efficiency catch up occurs at a constant rate over the benchmarking period, allowing no possibility for DNSPs to catch up more or less quickly over time;
- Quantonomics' concern over the apparent lack of convergence of some of the stochastic frontier analysis (SFA) models. We found that convergence could be achieved for all models if appropriate starting values are used; and
- Quantonomics' concern that efficiency scores could not be obtained for some DNSPs, in some models, using standard post-estimation commands in the statistical package Quantonomics used to estimate those models. We found that efficiency scores could be calculated for all DNSPs in all models using a different post-estimation command.

We investigated further variants of, and extensions to, the models considered by Quantonomics

In addition to the models investigated by Quantonomics, we also tested:

- versions of the BC95 and Kumb90 models that allow a different trend in efficiency for the Australian DNSPs, for each historical three-year period. This allowed the efficiency time trend in the BC95 and Kumb90 models to be estimated more flexibly;
- versions of the Kumb90 model that allowed the average efficiency adjustment factor for DNSPs in Australia, New Zealand and Ontario to differ; and
- versions of the time-varying LSE model that allow different efficiency time trends to be estimated for years before 2015 (the year in which the AER introduced the current benchmarking regime) and for the years post-2105. This allowed the models to capture any response by DNSPs to the introduction of formal benchmarking analysis within the DNSP regulatory framework.

We also tested translog versions of the existing SFA and LSE models used by the AER—as points of comparison against the time-varying inefficiency models we investigated.

Recommendations on models

Using model selection criteria very similar to those used by Quantonomics, we found that:

1. **The time-invariant inefficiency models currently used by the AER are fundamentally mis-specified and should be discontinued.** Their assumption of constant efficiency is inconsistent with observed data and leads to unreliable outcomes.

2. **Three time-varying SFA models satisfy all selection criteria and materially outperform all others tested:**

- SFA-BC95-JTT-HN-GTC

  This is a time-varying BC95 model that:

  - allows for different efficiency time trends for Australia, New Zealand and Ontario;
  - assumes a half-normal efficiency distribution in the final year of the sample; and
  - permits different cost time trend estimates for each three-year year historical period.

- SFA-BC95-AGEC-HN-GTC

  This is a time-varying BC95 model that:

  - allows different efficiency time trend estimates for each three-year historical period, for the Australian DNSPs, while allowing simple, linear efficiency time trends for the New Zealand and Ontarian DNSPs;
  - assumes a half-normal efficiency distribution in the final year of the sample; and
  - permits different cost time trend estimates for each three-year year historical period.

- SFA-Kumb-AGECJUR-HN-GTC

  This is a time-varying Kumb90 model that:

  - allows different efficiency time trend estimates for each three-year historical period, for the Australian DNSPs, while allowing simple, linear efficiency time trends for the New Zealand and Ontarian DNSPs;
  - allows different efficiency levels to be estimated for Australia, New Zealand and Ontario;
  - assumes a half-normal efficiency distribution; and
  - permits different cost time trend estimates for each three-year year historical period.

We recommend that these three models replace the existing (time-invariant) econometric models used by the AER. All three of these models should, by default, be given equal weight.

In addition, we concluded that:

3. **The LSE specifications should be abandoned.** These models are overly restrictive, and their statistical performance is unambiguously worse than that of the preferred SFA models.

4. **The AER should give primacy to translog specifications, using the Cobb-Douglas models only as a fallback if the translog models cannot be used due to serious statistical problems.** Formal statistical testing shows clearly that the Cobb-Douglas models fit the data

more poorly than the translog specifications. Therefore, these models are unnecessary provided the recommended translog models can be estimated reliably.

Recommendations on the process for estimating efficient base-year opex

The AER currently estimates the efficient level of opex in the base year by:

- using the estimate of the DNSP's *average* efficiency over the benchmarking period to derive an estimate of efficient opex in the middle of the benchmarking period; and then
- rolling that figure forward to the base year, accounting for output growth, technology change, and the change in undergrounding between the middle of the benchmarking period and the base year.

The AER could continue to use this process if it adopts time-varying inefficiency models in place of the existing benchmarking models. The estimate of average efficiency over the benchmarking period would be estimated by averaging the DNSP's annual efficiency scores over the period.

This would be the simplest approach, involving minimal change from the approach currently used by the AER, and understood by DNSPs and other stakeholders. This approach would also mean no change would be required to the way operating environment factor (OEF) adjustments are computed and applied in the benchmarking process.

Quantonomics has suggested that if time-varying inefficiency models are adopted by the AER, DNSPs' efficiency scores in the *final* year of the benchmarking period could be used to estimate efficient base year opex directly (if the base year corresponds to the final year of the benchmarking period), or by rolling forward from the final year (rather than the middle) of the period to the base year. The AER should be very cautious about adopting this approach for a number of reasons:

- The efficiency scores in individual years under the Kumb90 time-varying LSE models should not be interpreted as estimates of the DNSP's efficiency in those years, because of the deterministic way in which the inefficiency trend in those models are specified. Only the efficiency scores from (reliable) BC95 models should be interpreted in that way, because they are genuinely responsive to changes in DNSPs' opex from year to year;
- When rolling forward an estimate of efficient opex from the final year of the benchmarking period, the AER would need to do so in a way that does not disincentivise DNSPs from realising large efficiency gains in that year, so as to avoid base year adjustments and a removal of Efficiency Benefit Sharing Scheme (EBSS) rewards; and
- The AER would need to modify the way it calculates OEF adjustments to ensure that those adjustments reflect cost differences between the DNSP in question and the reference DNSPs in the final year of the benchmarking period, rather than the average difference in costs over the benchmarking period.

On balance, given these complications, we recommend that the AER retain its existing opex roll-forward approach and only make changes to this approach following proper consultation with stakeholders.

1.3 Structure of this report

The remainder of this report is structured as follows:

- Section 2 provides an overview of the Phase 2 report and Quantonomics' key findings;
- Section 3 explains a number of issues we identified with the analysis in the Phase 2 report that needed to be addressed before the models in that report could be evaluated properly;
- Section 4 introduces the models we have investigated in this report, including extensions and modifications of the models tested by Quantonomics;
- Section 5 presents our evaluation of the models we tested; and
- Section 6 discusses the issues involved with estimating efficient base year opex under time-varying inefficiency models.

2 Overview of Quantonomics Phase 2 report

2.1 Key models considered by Quantonomics

Quantonomics' investigations focussed on several different classes of time-varying inefficiency models, of which three receive the most attention:

- the Battese and Coelli (1995) (BC95) model (including a demonstration model we presented in our response to the Phase 1 consultation for purely illustrative purposes);
- the Kumbhakar (1990) (Kumb90) model;
- the least squares econometrics (LSE) model;
- the Colombi et al (2014) (Four Components) model; and
- the Cornwell et al (1990) (FECSS) model.

Of these five classes of models, most attention in Quantonomics' Phase 2 report is devoted to the assessment of the BC95, Kumb90 and LSE models.

Quantonomics found that the estimated efficiency scores from the Four Components model were unreliable (nearly all the estimated scores were close to 1), most likely because the model was unable to disentangle well the effects of inefficiency and heterogeneity between DNSPs. For this reason, the Four Components model was rejected by Quantonomics.

Quantonomics found that the estimate output elasticities from the FECSS model indicated poor model fit, with extremely high monotonicity violations and statistically insignificant output coefficients. For these reasons, the FECSS model was rejected by Quantonomics.

Sections 2 and 4 of Quantonomics' Phase 2 report provides a clear and comprehensive exposition of the BC95, Kumb90 and LSE models. Hence, for conciseness, we do not repeat that detail here. However, the remainder of this section provides a brief summary of the key features of each of these models, and how they were implemented by Quantonomics—including the treatment of DNSP and jurisdictional time trends.

2.1.1 BC95 models

Comparison to existing AER SFA model

Before discussing BC95 models, it is useful to briefly review AER's existing stochastic frontier analysis (SFA) model since the BC95 model is an extension of the AER's existing SFA model. Let the opex cost function be $c_{it} = \alpha_0 + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \epsilon_{it}$, where c_{it} is the log of the measure of cost for firm i in period t, α_0 is the intercept, $\boldsymbol{x}_{it}$ is a vector of cost drivers, $\boldsymbol{\beta}$ is a vector of slope parameters (elasticities), and ϵ_{it} is the error term.

The AER's standard model follows the Pitt and Lee (1981) approach,[5] which assumes that the error term consists of two unobserved components: cost inefficiency (u_i) and statistical noise (v_{it}). Specifically, it assumes $\epsilon_{it} = u_i + v_{it}$, with a restriction that $u_i \geq 0$, reflecting that inefficiency can only be non-negative. The inefficiency u_i is interpreted as a random effect, following a parametric distribution. Each DNSP is assumed to draw randomly from a common efficiency distribution once and always maintains that level of inefficiency. This is what gives rise to the time-invariant inefficiency of the DNSPs in the AER's existing SFA model.

This random effect model cannot be estimated using a closed-form solution; it must be estimated using maximum likelihood estimation, which requires numerical optimisation and convergence.

The AER's standard model assumes that $u_i \geq 0$ follows a truncated normal distribution, or $u_i \sim N^+(\mu, \sigma_u^2)$ where μ is the pre-truncation mean. The estimated inefficiency can be obtained from $\hat{u}_i = E(u_i|\epsilon_{it})$.

The BC95 model allows inefficiency $u_{it} \geq 0$ to vary over time and across DNSPs. The model assumes $u_{it} \sim N^+(\mu_{it}, \sigma_u^2)$ with $\mu_{it} = \delta_0 + \boldsymbol{z}_{it}'\boldsymbol{\delta}$, where δ_0 is the intercept of the pre-truncation mean, $\boldsymbol{z}_{it}$ is a vector of drivers of the pre-truncation mean, and $\boldsymbol{\delta}$ is the vector of slope parameters. Under the BC95 model, each DNSP is assumed to draw randomly from a common efficiency distribution each year, thus allowing inefficiency to vary over time.

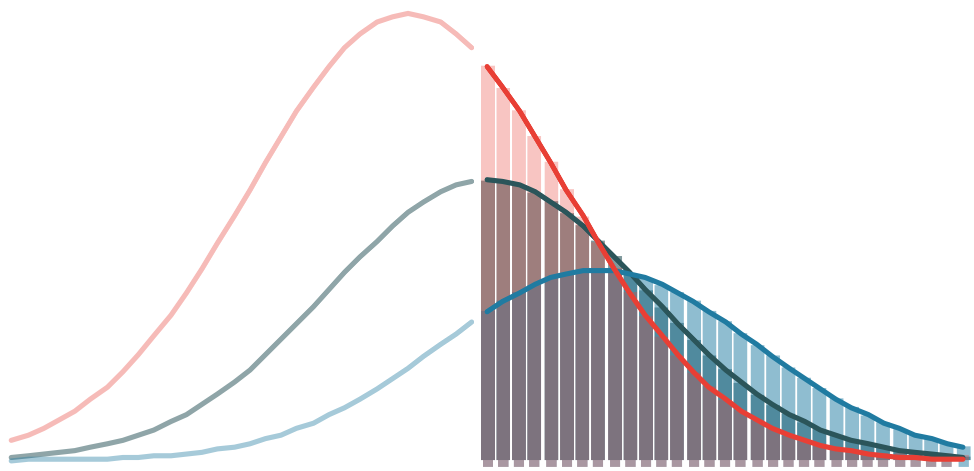
The mode of the distribution is estimated via maximum likelihood estimation. The estimated inefficiency can be obtained from $\hat{u}_{it} = E(u_{it}|\epsilon_{it})$.[6]

Figure 1 illustrates the inefficiency distributions in the BC95 models. The figure shows three truncated normal distributions, each having a different pre-truncation mean. The black curve assumes $\mu_{it} = 0$ (half-normal). The red curve assumes $\mu_{it} < 0$. The blue curve assumes $\mu_{it} > 0$. The higher the μ_{it}, the more dispersed is the inefficiency distribution.

5 Pitt M and Lee L (1981), The measurement and sources of technical inefficiency in the Indonesian weaving industry, Journal of Development Economics, 9(1).

6 In practice Quantonomics estimates efficiency using $E(\exp(-u_{it})|\epsilon_{it})$. This is consistent with the approach in Battese and Coelli (1988). See Battese G and Coelli T (1988), Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data, Journal of Econometrics, 38(3).

Figure 1: Inefficiency distributions in BC95 models



Note: BC95 models assume inefficiency $u_{it} \geq 0$ follows a truncated normal distribution, or $u_{it} \sim N^+(\mu_{it}, \sigma_u^2)$ where μ_{it} is the pre-truncation mean of a normal distribution. The figure shows three truncated normal distributions, each having a different pre-truncation mean. The black curve assumes $\mu_{it} = 0$ (half-normal). The red curve assumes $\mu_{it} < 0$. The blue curve assumes $\mu_{it} > 0$. The higher the μ_{it}, the more dispersed is the inefficiency distribution.

BC95 models can capture time-varying inefficiency flexibly. Arguably the simplest specification is one in which the pre-truncation mean follows a linear time trend: $\mu_{it} = \delta_0 + \delta t$ where $\delta < 0$ implies reducing inefficiency over time. In the figure, this suggests the inefficiency distribution shifts from the blue curve to the black curve, and then to the red curve. In this specification, firms improve their efficiencies on average over time, and they tend to become more similar in efficiency and converge to being fully efficient over time.

Quantonomics implementation

Quantonomics assumes that the intercept of the pre-truncation mean is zero, or $\delta_0 = 0$. This is referred to as the "no constant term" specification for μ_{it}. Quantonomics' justification for this normalisation is that it is difficult to separately identify δ_0 from the cost/frontier intercept α_0.[7] Quantonomics reported that setting $\delta_0 = 0$ could mitigate numerical convergence problems that they encountered.[8]

Quantonomics considers μ_{it} to be determined by juristicational or DNSP-specific time trends, with no constant term:

- Jurisdictional time trends (JTT): Each jurisdiction (Australia, New Zealand and Ontario) has its own linear time trend in efficiency.

7 Quantonomics Phase 2 report, pp. 18-19.

8 Although Quantonomics did not point this out, setting $\delta_0 = 0$ implies that the pre-truncation mean $\mu_{it} = 0$ when $\boldsymbol{z}_{it} = 0$. This in turn implies that $u_{it} \sim N^+(0, \sigma_u^2)$, or u_{it} follows a half-normal distribution when $\boldsymbol{z}_{it} = 0$.

- Australian DNSP-specific time trends and jurisdiction-specific time trends (AJTT): Each Australian DNSP has its own linear time trend in efficiency. Under this specification, all New Zealand DNSPs are assumed to have a common New Zealand-specific linear time trend in efficiency and all Ontarian DNSPs are assumed to have a common (but separate from New Zealand) Ontario-specific linear time trend in efficiency.

Quantonomics also consider two types of cost/frontier trends: Linear time trend and a GTC specification, which is a step change in cost every three years.

This yields four main models: JTT/AJTT in efficiency trend combined with linear/GTC in cost trend.

Quantonomics found that the AJTT variants of the BC95 model is more difficult to estimate due to its complexity. Hence, Quantonomics' preferred variant is JTT efficiency trend with linear cost trend.

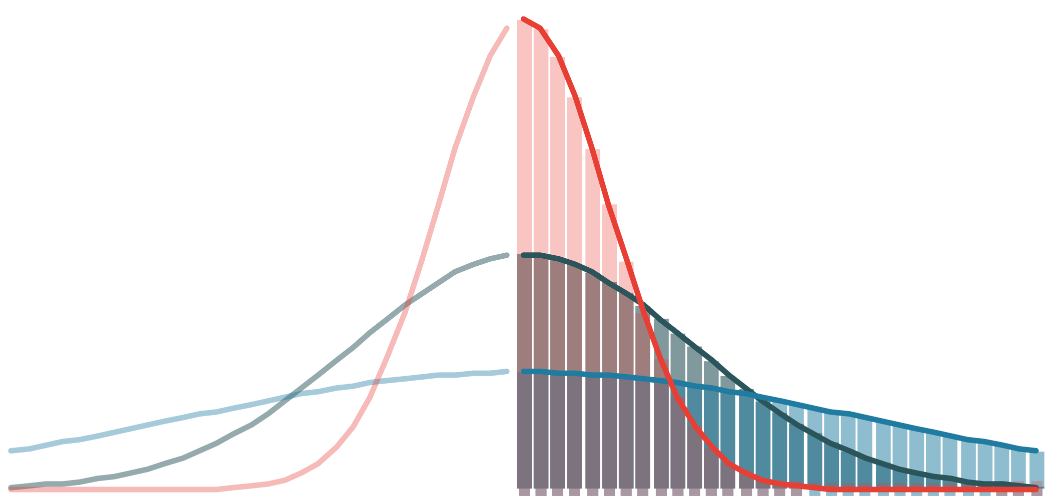
2.1.2 Kumb90 models

Overview of the model

The Kumb90 model is also an SFA model. It allows for time-varying inefficiency via a deterministic, multiplicative scaling factor applied to the inefficiency term u_i. Specifically, Kumb90 assumes $u_{it} = g(\boldsymbol{z_{it}})u_i \geq 0$, where the scaling factor $g(\boldsymbol{z_{it}}) \geq 0$ is a deterministic function of a vector of drivers $\boldsymbol{z_{it}}$, and $u_i \geq 0$ follows a half-normal distribution, or $u_i \sim N^+(0, \sigma_u^2)$. Like the standard SFA and BC95 models, the Kumb90 model must be estimated numerically, rather than using a closed-form solution, using maximum likelihood estimation.

Figure 2 below illustrates the inefficiency distributions in Kumb90 models. The figure shows three half-normal distributions, each having a different scale. The black curve assumes $g(z_{it}) = 1$. The red curve assumes $g(z_{it}) < 1$. The blue curve assumes $g(z_{it}) > 1$. The higher the scaling factor, the more dispersed is the inefficiency distribution.

Figure 2: Inefficiency distributions in Kumb90 models



Source: Kumb90 models assume inefficiency $u_{it} = g(z_{it})u_i \geq 0$ follows a scaled half-normal distribution, or $u_i \sim N^+(0, \sigma_u^2)$ and $g(z_{it})$ is a scaling factor. The figure shows three half-normal distributions, each having a different scale. The black curve assumes $g(z_{it}) = 1$. The red curve assumes $g(z_{it}) < 1$. The blue curve assumes $g(z_{it}) > 1$. The higher the scaling factor, the more dispersed is the inefficiency distribution.

The scaling factor $g(\boldsymbol{z_{it}})$ must be non-negative. The original version of the Kumb90 model assumes $g(\boldsymbol{z_{it}}) = \frac{1}{1+\exp(at+bt^2)}$, which is a logistic function (sigmoid curve) bounded between 0 and 1. This specification assumes the latent inefficiency index, $at + bt^2$, to exhibit a quadratic time trend. If $a > 0$ and $b = 0$, the index becomes linear and increasing in t, and the scaling factor $g(z_{it})$ is an inverse S-shape curve shrinking toward 0 over time. Firms improve their efficiencies on average over time, and they tend to become more similar in efficiency and converge to being fully efficient over time.

In the specification above, even though the index is linear in t, the mean of inefficiency $u_{it} = g(\boldsymbol{z_{it}})u_i$ is a nonlinear function of time, that is, the rate of efficiency change can become faster or slower over time. More generally, the specification of $g(\boldsymbol{z_{it}})$ will determine how the mean and variance of the inefficiency distribution will change, nonlinearly, over time.

The *sfpanel* implementation command for Kumb90 assumes $g(\boldsymbol{z_{it}}) = \frac{1}{1+\exp(\delta_0 + \boldsymbol{z}'_{it}\boldsymbol{\delta})}$, where δ_0 is the intercept of the latent index and $\boldsymbol{\delta}$ is a vector of slope parameters.

Kumb90 models only allow for more restrictive forms of time-varying inefficiency. Unlike the BC95 models, each firm only draws u_i once, and then its inefficiency $u_{it} = g(\boldsymbol{z_{it}})u_i$ evolves over time deterministically via scaling factor $g(\boldsymbol{z_{it}})$. More specifically, suppose $z_{it} = t$, so that $g(\boldsymbol{z_{it}}) = \frac{1}{1+\exp(\delta_0+\delta t)}$. Then, inefficiency u_{it} evolves over time proportionally by the realised magnitude of u_i. This means that firms with larger draws of u_i (i.e., those that are less efficient) will have a faster change in efficiency over time. Quantonomics calls this a "common trend" restriction, namely, the inefficiencies of firms differ in magnitude but follow the same time pattern.

Quantonomics implementation

Quantonomics considers the index $\delta_0 + \boldsymbol{z}'_{it}\boldsymbol{\delta}$ to be determined by juristicational or DNSP-specific time trends:

- Jurisdictional time trends (JTT): Each jurisdiction (Australia, New Zealand and Ontario) has its own linear time trend in efficiency.
- Australian DNSP-specific time trends and jurisdiction-specific time trends (AJTT): Each Australian DNSP has its own linear time trend in efficiency. Under this specification, all New Zealand DNSPs are assumed to have a common New Zealand-specific linear time trend in efficiency and all Ontarian DNSPs are assumed to have a common (but separate from New Zealand) Ontario-specific linear time trend in efficiency.

Quantonomics also consider two types of cost/frontier trends: Linear time trend and a GTC specification, which is a step change in cost every three years.

This yields four main models: JTT/AJTT in efficiency trend combined with linear/GTC in cost trend.

Quantonomics found that the AJTT variants of the BC95 model is more difficult to estimate due to its complexity. Hence, Quantonomics' preferred variant is JTT efficiency trend with linear cost trend.

2.1.3 LSE models

Standard LSE model

Unlike the BC95 and Kumb90 models, the LSE model is not an SFA model, in which the aggregate error term is decomposed into an inefficiency term and noise for all DNSPs. Instead, the LSE model implements a fixed-effects panel data regression, estimating the fixed effects for Australian DNSPs as cost shifters, and using the relative costs shifters to construct efficiency estimates for Australian DNSPs.[9] The relative magnitudes of fixed effects represent the DNSPs' permanent differences in cost.

A key assumption of the LSE model is that at least one DNSP is at the frontier and achieves 100% efficiency. The inefficiencies of other firms are then computed as the difference in fixed effects between them and that of the frontier DNSP.[10]

The AER already uses a version of the LSE model, which assumes DNSP inefficiency is time-invariant. Specifically, the AER currently implements the LSE model using firm-specific dummy variables to estimate the fixed effects, but only for Australian DNSPs. It uses jurisdictional dummy variables (one for New Zealand DNSPs and another for Ontarian DNSPs) to capture differences in the frontier between Australia and other jurisdictions. It uses an aggregate linear time trend in cost to capture the change in frontier of real opex over time.

Quantonomics implementation

Quantonomics examined alternative methods of estimating the LSE models, in particular, ones that allow for time-varying efficiency. It extended standard implementation by incorporating one of the following:

- Australian DNSP-specific time trends (ADTT): Each Australian DNSP has its own linear time trend in cost.
- Australian DNSP-specific time trends and jurisdiction-specific time trends (AJTT): This is ADTT plus all New Zealand DNSPs having a common linear time trend in cost. This in effect allows New Zealand DNSPs, collectively, to have time varying efficiency, while implicitly Ontario DNSPs are assumed to have time-invariant efficiency.

In both approaches, Quantonomics assumes that in each year at least one Australian DNSP is at the efficient frontier (i.e., that DNSP is assumed to be 100% efficient). In each year, Quantonomics computes, for each Australian DNSP, a composite term $\hat{\alpha}_i + \hat{\delta}_i t$, where $\hat{\alpha}_i$ is the DNSP's fixed effect and $\hat{\delta}_i$ is the DNSP's time trend coefficient.[11] It then obtained each Australian DNSP's efficiency as $\hat{u}_{it} = \hat{\alpha}_i + \hat{\delta}_i t - \min_{j \in AusDNSPs}\{\hat{\alpha}_j + \hat{\delta}_j t\}$. The difference between this and the standard LSE model's computation is the time trend coefficient $\hat{\delta}_i$, which is allowed to be non-zero.

Relaxing the aggregate linear trend in cost, Quantonomics also examines a GTC specification, which is a step change in cost every three years.

9 This approach is in line with Schmidt and Sickles (1984). The AER's existing LSE models are estimated using the Stata command *xtpcse*, which accommodates autocorrelation of errors within the panel and produces panel-corrected standard errors.

10 Such an assumption is required as no assumption is made regarding the distribution of efficiency, unlike the SFA models.

11 With $\hat{\delta}_i$ equal for all Australian DNSPs under the ADTT specification.

Combining the ADTT/AJTT specifications in efficiency and linear/GTC specifications in cost, Quantonomics examined four LSE model specifications:

- LSE-ADTT;
- LSE-ADTT–GTC;
- LSE-AJTT; and
- LSE-AJTT-GTC.

Quantonomics argued that a key advantage of LSE models is the simplicity of estimation, relative to SFA models. LSE models do not have the same computational complexity as SFA models, where convergence to a global optimum is not guaranteed and is reliant on selection of appropriate starting values. With LSE models there is no convergence issue, and no starting values are required as the estimation provides a "closed form" solution rather than iteratively searching for better-fitting parameter estimates.

Quantonomics concluded that at least two of its LSE models were among the four best-performing models. It recommended an LSE model to be included alongside SFA models to improve the reliability of the estimated efficiency scores.

2.1.4 Other models considered

Quantonomics explored a few other time-varying inefficiency models.[12] This included estimating an additional two specifications:

- The Colombi et al (2014) Four Components model; and
- The Cornwell et al (1990) FECSS[13] model.

Under the Four Components approach, the distance to the efficient frontier (i.e., the regression residual) is separated into four distinct components: time-invariant inefficiency, time-varying inefficiency, firm specific heterogeneity and random noise. By contrast, the standard SFA model seeks to decompose the regression error term into just two components: inefficiency and random noise.

While Quantonomics considered the Four Components to be theoretically appealing,[14] it concluded that the results from the model did not perform well in practice, noting that:

- The estimated coefficient on the customer numbers output variable was not statistically significant for the Cobb-Douglas and translog short sample models;
- Both the short-sample and long-sample models suffered from a large number of monotonicity violations; and
- The average estimated efficiency scores for all DNSPs were implausibly high (e.g., over 99% for the short-sample models).

12 Quantonomics Phase 2 report, pp. 8-20.

13 Fixed Effects Cornwell, Schmidt, and Sickles.

14 Quantonomics Phase 2 report, p. 44.

For these reasons, Quantonomics concluded that the Four Components model should not be considered further.

The FECSS model estimates a fixed-effect efficiency term for each DNSP in the sample and then applies an adjustment to obtain time-varying efficiency, with the adjustment estimated as a DNSP-specific quadratic function of time. This approach shares some similarities with the Kumb90 model, in that a time varying adjustment is applied to the time invariant component. Quantonomics found that the FECSS models similarly failed to perform well, noting that:[15]

- Almost none of the estimated coefficients associated with the output variables were found to be statistically insignificant under the four models (i.e., the Cobb-Douglas and translog models over the short and long benchmarking periods) estimated;
- The models suffered from a large number of monotonicity violations.

This poor performance of the FECSS models was attributed to over-fitting. Fixed effects were estimated for each DNSP (including the New Zealand and Ontarian DNSPs) which removes the ability for variation across DNSPs to inform estimates of the cost function (the cost function is estimated based on how changes in outputs over time lead to changes in costs for DNSPs). In addition, many efficiency trend variables needed to be estimated, with two (a linear trend and quadratic trend) for each DNSP. Given the large number of parameters that needed to be estimated, Quantonomics considered that the FECSS models would be impractical for the AER's purposes.

Given that Quantonomics rejected both the Four Components model and the FECSS models, the Quantonomics analysis focussed on the BC95, Kumb90 and LSE models.

2.2 Key variations to the models explored by Quantonomics

2.2.1 Half-normal assumption

The AER's existing SFA models feature an inefficiency term follows a truncated normal distribution—i.e., a normal distribution with mean μ and variance σ^2 that is truncated at zero, with only the positive portion of the distribution used. Previously this has resulted in an identification issue, where it is challenging to distinguish between parameter estimates with high efficiencies but expensive cost functions (low μ and a high cost function intercept term) and parameter estimates with low efficiencies but less expensive cost functions (high μ and a low cost function intercept term). This may have contributed to previous challenges in estimating the SFA models.

To address this, Quantonomics proposes to restrict the mean of the non-truncated inefficiency distribution to zero, such that the truncated normal is in fact a half-normal distribution (see, for example, the black curve in Figure 1 and all three curves in Figure 2). This simplifies the estimation of the models considerably and is found to ameliorate the convergence issues.[16]

Although the assumption of a half-normal distribution for the inefficiency term in SFA models is more restrictive than the original truncated normal distribution, it is not unreasonable given the improvement in the convergence of the models.

[15] With a single exception.

[16] Quantonomics Phase 2 report, p. 25.

2.2.2 GTC time trend

To relax the linear time trend assumption, Quantonomics adopted a 'general technical change' (GTC) approach, which follows Baltagi and Griffin (1988).[17] The full sample period is divided into six non-overlapping periods: 2006-2008; 2009-2011; 2012-2014; 2015-2017; 2018-2020; 2021-2023. Each of these periods, except for 2006-2008 which is used as the reference period (excluded category), is represented by a dummy variable. This results in five dummy variables.[18] The coefficients of these dummy variables represent the step change in time trend relative to 2006-2008. Sets of years in 3-year blocks are used rather than individual years to conserve degrees of freedom.

Quantonomics only considered GTC time trends in cost/frontier. They assumed linear efficiency trends, which is a strong assumption. In contrast, we considered GTC time trends in cost/frontier and in efficiency. This allows the rate of efficiency change to become faster or slower, or even change direction, over time.

We think that this is, in principle, a good approach. As explained in section 4, we extend the concept of the GTC time trend to model a 'generalised efficiency change' (GEC)—a time trend associated with the efficiency of Australian DNSPs.

2.2.3 Heteroscedasticity

Quantonomics considered heteroscedasticity specifications of the BC95 models. The BC95 approach, as implemented in the *sfpanel* package in Stata, can allow the variance of the idiosyncratic noise term v_{it} to vary across observations. In its analysis, Quantonomics considers whether the scale (i.e., size) of the DNSP could impact the variance of the noise term.

To examine this, Quantonomics found the mean of (log) circuit length, customer numbers and ratcheted maximum demand for each DNSP. It then weighted these three output measures to obtain a measure of the average size of the DNSP.[19]

Quantonomics did not find this approach to be successful, consistent with weak relationships between residuals and size. Consequently, we do not consider these models in our own analysis.[20]

2.2.4 Consideration of omitted variables

While the standard opex cost function only includes three outputs, Quantonomics noted that the total factor productivity (TFP) index analysis includes an additional two outputs: (energy delivered, and customer minutes off-supply). Quantonomics therefore investigated whether including these two additional output variables in the econometric models would improve them, and whether their omission would bias parameter estimates of the cost function.

Quantonomics tested each of the standard models with the addition of energy delivered, customer minutes off-supply (CMOS) and both. No improvement was observed with the addition of energy delivered, and the models often worsened considerably in terms of monotonicity violations.

17 Baltagi BH and Griffin JM (1988), A General Index of Technical Change, Journal of Political Economy 96(1).

18 Additional dummy variables can be introduced to the model as additional years are added.

19 Weights obtained from a simple regression including the Cobb-Douglas specification, a time trend and jurisdiction dummies.

20 Quantonomics Phase 2 report, p. 50.

Quantonomics noted that the inclusion of CMOS may constitute a very slight improvement over the standard models, but considered that CMOS as an explanatory variable could introduce an endogeneity problem that could result in biased model estimates.

For models that included both energy delivered and CMOS as explanatory variables, similar issues were observed, including a large number of monotonicity violations and no model improvement. Quantonomics therefore concluded that it would be "impractical" to include energy delivered and CMOS as additional output variables in the econometric benchmarking models.

Quantonomics also explored the inclusion of input prices into the models, noting that input prices are a typical component of cost functions. After testing, Quantonomics found that the input prices may also be an omitted variable but, similar to CMOS, the inclusion of input prices did not materially improve the model in either goodness of fit or monotonicity violations. Quantonomics also noted that there is uncertainty over the measurement of the price of capital inputs under the AER's user cost of capital approach, which would need to be resolved before capital input prices could be incorporated reliably into the models. On these grounds, Quantonomics decided not to incorporate input prices as explanatory variables into the models.

2.2.5 Summary of main models explored by Quantonomics

Table 1 below summarises the variations of the BC95, Kumb90 and LSE models investigated by Quantonomics.

Table 1: **Key models investigated by Quantonomics**

BC95		Kumb90		LSE	
Efficiency trend	**Cost trend**	**Efficiency trend**	**Cost trend**	**Efficiency trend**	**Cost trend**
JTT: Aus, NZ, Ont (linear)	Linear	**JTT**: Aus, NZ, Ont (linear)	Linear	**ADTT**: AusDNSP (linear)	Linear
	GTC		GTC		GTC
AJTT: AusDNSP, NZ, Ont (linear)	Linear	**AJTT:** AusDNSP, NZ, Ont (linear)	Linear	**AJTT+NZ**: AusDNSP, NZ (linear)	Linear
	GTC		GTC		GTC

- GTC: step change in cost function every 3 years.
- JTT: jurisdictional time trend
- AJTT: Aus DNSP-specific & jurisdictional time trend
- ADTT: Aus DNSP-specific time trend
- ADTT+NZ: Aus DNSP specific & NZ time trend

Note:

- *SFA models allow all firms to be inefficient.*
- *LSE models need reference; at any point in time, some firms are right at frontier.*

Source: Frontier Economics.

2.3 Assessment criteria

This section summarises the assessment criteria applied by Quantonomics.

2.3.1 Sign and significance of (primary) output coefficients

Quantonomics examined the sign and statistical significance of the three primary output variables: circuit length, customer numbers and ratcheted maximum demand.[21] These output variables are considered by the AER to be important drivers, and so negative estimated coefficients, or coefficient estimates that are not statistically significant, for these outputs may indicate that the model is unreliable.

The results for Quantonomics' models are presented in Tables 5.2 and 5.3 of the Quantonomics report.[22]

2.3.2 Frequency of monotonicity violations in the translog models

Quantonomics considered the frequency of monotonicity violations in the translog models among Australian DNSPs.

Quantonomics compared the results of time-varying models to the corresponding standard models, with fewer monotonicity violations taken to mean better performance. We note that the standard SFA model has severe issues: monotonicity violations are presented for 79.5% of observations for the long sample translog model. Similarly, the long sample LSE translog model suffers from a large number of monotonicity violations (22.2% of the Australian sample).

The results for Quantonomics' models are presented in Tables 5.2 and 5.3 of the Quantonomics report.[23]

2.3.3 Reasonableness of output weights

The coefficients on the primary output variables are considered, both the size of the individual estimates but also the sum (i.e., the degree of economies of scale).

While Quantonomics cites comparisons with index weights and prior econometric results for this criterion,[24] in practice analysis was focused on comparisons across the estimated models.[25]

2.3.4 Coefficients of other variables have the correct sign

While not considered an output variable, the underground share is considered a driver of opex, with higher undergrounding associated with lower opex.

21 We note that, for translog models, the coefficients represent the elasticity of cost with respect to outputs at the sample average outputs.

22 Quantonomics Phase 2 report, pp. 82-83.

23 Quantonomics Phase 2 report, pp. 82-83.

24 Quantonomics Phase 2 report, p. 22.

25 Quantonomics Phase 2 report, pp. 84-86.

Quantonomics examines the sign on undergrounding, and in circumstance where it is the opposite to expectations it examines the significance. The results for Quantonomics models are presented in Tables 5.2 and 5.3 of the Quantonomics report.[26]

2.3.5 Convergence

SFA models may not solve correctly: in some cases the model may fail to converge. If this occurs the model results are disregarded as no valid results have been estimated. To a lesser extent this includes models that result in missing standard errors.

2.3.6 Goodness-of-fit

Quantonomics evaluated goodness of fit on the basis of comparing measures like the Bayesian Information Criterion (BIC) for models sufficiently similar to each other, as well as other measures like the adjusted R-squared and the pseudo-adjusted R-squared where appropriate. We note that these two criteria favour parsimony by penalising more complicated models (i.e., those with more parameters that require estimation). Quantonomics prefers models with higher goodness of fit statistics, *"provided they also satisfy theoretical and specification criteria"*. [27]

2.3.7 Performance on specification tests

Quantonomics performed specification tests primarily using the *linktest* command in Stata. This command regresses the observed values (in this case, real historical opex) against:

- the predicted/fitted values from the model being tested; and
- the square of those fitted values.

When performing this regression, the *linktest* command allows the user to specify the structure of the model being tested. If, for example, the model being tested is the SFA-BC95-JTT-HN-GTC model, and the user specifies this, then the *linktest* regression will impose the following structure:

- The model is the BC95 model with linear efficiency trend (in the way defined by Quantonomics);
- The model allows for different jurisdictional efficiency time trends for Australia, New Zealand and Ontario; and
- Inefficiency is assumed to be half-normally distributed in the final year of the sample.

If the *square* of the fitted values is significant, that is an indication that the model is mis-specified such that the square of the fitted values explains some of the structure of the data. Where *linktest* was not available due to model incompatibility with the function, Quantonomics referred to residuals plots to identify any remaining structure in the dataset that the model does not address. Quantonomics also utilised the *collin* command to examine multicollinearity effects.[28]

26 Quantonomics Phase 2 report, pp. 82-83.

27 Quantonomics Phase 2 report, pp. 23.

28 Quantonomics Phase 2 report, pp. 23.

2.3.8 Stability to sample changes

Quantonomics considered the extent to which parameter estimates were stable against changes in the underlying data. This was performed by using subsets of the 2006-23 data (increasing the start year or decreasing the end year) and examining changes in results. This was performed for a subset of models.[29] The performance of the models to sample changes was examined by considering how well the models converged, coefficient estimates and the extent of monotonicity violations: key indicators of model issues.

2.3.9 Consistent efficiency scores

Quantonomics considered the reasonableness of efficiency scores produced by different models. Consistency was evaluated with respect to other econometric models and Opex MPFP results, examining both the time pattern of efficiency score estimates and the overall level of efficiency.

This was primarily examined using correlations with Opex MPFP scores,[30] though the level of efficiency is considered and used to exclude models that produced implausibly high or low scores.[31]

2.3.10 Parsimony

Quantonomics considered the extent to which models were overly complex. All else equal, simpler model were preferred over more complex models that may be more challenging to implement or interpret.

2.4 Findings

Quantonomics assessed the performance of the models against various selection criteria and identified four models that it considered perform best according to those criteria:

- one BC95 model (i.e., BC95-JTT-HN);
- one Kumb90 model (i.e., Kumb90-JTT-HN); and
- two LSE models (i.e., LSE-AJTT and LSE-AJTT-GTC).

Quantonomics stated that:

> *Overall, these models produce results that are consistent with expected output elasticities and the correct signs of coefficients, particularly in the longer-period sample. They also exhibit few or no monotonicity violations, a high level of goodness of fit, satisfactory performance in statistical tests, and stable efficiency scores.*[32]

29 Quantonomics Phase 2 report, pp. 110-115.

30 Quantonomics Phase 2 report, pp. 105-108.

31 See the analysis of the Colombi models, Quantonomics Phase 2 report, pp. 45

32 Quantonomics Phase 2 report, p. 110.

However, Quantonomics stopped short of recommending that these models should replace the AER's existing econometric models, even though Quantonomics' analysis shows they objectively outperform the existing models. Quantonomics stated that:

> *...while these findings highlight the potential value of incorporating time-varying inefficiency models into the opex assessment framework, they remain preliminary. We reiterate that these are initial views and further analysis is required.*[33]

2.5 Use of final year efficiency for benchmarking

Historically, the AER has estimated DNSPs' efficient base year opex using the benchmarking roll forward model (BRFM) approach.[34] Under this approach, the sample average real opex of the DNSP is adjusted down to account for any difference between the target efficiency score (i.e. 75% adjusting for OEFs) and the period average efficiency score estimated using the econometric benchmarking models:

$$C_{PA}^{e} = C_{PA} \times \min\left\{\frac{\theta_{PA}}{75\%/(1 + OEF\%)}, 1\right\}$$

Where C_{PA} is the real period-average opex, C_{PA}^{e} is the efficient period-average opex, θ_{PA} is the period-average estimated efficiency score, and $OEF\%$ is the average adjustment for differences in operating environment factors (OEFs) between the DNSP in question and the reference DNSPs.

This provides an estimate of the average opex over the period if the DNSP were not materially inefficient (i.e., the DNSP's average efficient opex over the benchmarking period). The BRFM then rolls forward that opex estimate to the base year, accounting for output growth, technology change, and the change in undergrounding between the period average and the base year used to forecast efficient opex over the next regulatory period.

Quantonomics notes that the time-varying models can "enable the efficient base-year opex to be calculated directly if the base year is the same as the last year of the sample period."[35] Quantonomics suggests that "This can remove the need to rely on the benchmarking roll-forward model (BRFM) framework or reduce its role to bridging between the last year of the sample and the base year."[36] However, Quantonomics notes that "The BRFM can still be applied to the period-average efficiency scores produced by time-varying inefficiency models."[37]

In section 6.3 of its Phase 2 report, Quantonomics compares FY2023 efficiency estimates from various time-varying models against an alternative, called the "BFRM-implied efficiency scores for 2023", which is calculated as the ratio between (1) the efficient FY2023 opex obtained from applying the BRFM and (2) the actual FY2023 opex.

33 Quantonomics Phase 2 report, p. 143.

34 For DNSPs other than those identified as reference (i.e., efficient) DNSPs.

35 Quantonomics Phase 2 report, p. v.

36 Quantonomics Phase 2 report, p. v.

37 Quantonomics Phase 2 report, p. v.

A priori, one would expect the two measures to be highly correlated, as both are seeking to measure the efficiency of a DNSP in FY2023 opex. While Quantonomics notes that the correlation is only moderate for the Kumb90-JTT-HN model, Quantonomics concludes that the results of Table 6.4 "provide additional confidence in the reliability of the efficiency estimates from the selected models."[38]

Quantonomics concluded in Section 6.4, noting its findings "highlight the potential value of incorporating time-varying inefficiency models into the opex assessment framework", while noting that findings are preliminary and "further analysis is required."[39]

In Section 6 of this report we consider various options to estimate efficient base year opex, including the current BRFM (using the sample average time varying efficiency) in addition to adopting an alternative approach using the actual opex and the efficiency estimate in the final year of the sample to either:

- Directly estimate the efficient base year opex if the base year is the final year of the sample; or
- Use a modified BRFM to bridge from the final year of the sample to the base year if the base year occurs after the final year of the sample.

[38] Quantonomics Phase 2 report, p.142.

[39] Quantonomics Phase 2 report, p.143.

3 Issues identified in the Phase 2 report

When reviewing Quantonomics' Phase 2 report, we identified a number of issues with the Quantonomics' analysis that needed to be addressed before the time-varying inefficiency models could be assessed properly. These issues included:

- Quantonomics' treatment of the efficiency time trend in some BC95 models that results in estimates of expected efficiency that are essentially time invariant;
- The restrictive nature of the efficiency time trend in the Kumb90 models that tend to produce highly implausible estimates of efficiency over time;
- The restrictive assumption in LSE models that efficiency catch up occurs at a constant rate over the benchmarking period, allowing no possibility for DNSPs to catch up more or less quickly over time;
- Quantonomics' concern over the apparent lack of convergence of some of the stochastic frontier analysis (SFA) models; and
- Quantonomics' concern that efficiency scores could not be obtained for some DNSPs, in some models, using standard post-estimation commands in the statistical package Quantonomics used to estimate those models.

This section explains each of these issues and how we addressed them before evaluating the time-varying inefficiency models.

This section concludes by discussing briefly one set of possible omitted variables that were not investigated by Quantonomics, but that could have a material impact on benchmarking outcomes going forward.

3.1 Treatment of time trend in the efficiency component

Explanation of the issue

The manner in which efficiency is assumed to be time-varying is a key consideration in the models developed by Quantonomics. In Quantonomics' suite of eight BC95 models, the distribution of efficiencies u_{it} is allowed to vary over time. More specifically, the mean of the pre-truncation inefficiency term , μ_{it}, varies over time linearly, with $\mu_{it} = \beta_1 \times aus\_yr_{it} + \beta_2 \times nz\_yr_{it} + \beta_3 \times ont\_yr_{it}$, where aus_yr_{it} measures time for Australian firms, nz_yr_{it} measures time for NZ firms and ont_yr_{it} measures time for Ontario firms, and β_1, β_2 and β_3 reflect the impact of time on inefficiency in Australia, New Zealand and Ontario DNSPs, respectively.

Inspection of Quantonomics' Stata .do files shows that for the BC95-JTT-HN model:[40]

- aus_yr_{it} is equal to the year (i.e., 2006 to 2023) if DNSP i is in Australia, and zero otherwise;

[40] A similar approach was applied for the BC95-AJTT-HN and BC95-AJTT-HN-hY models.

- nz_yr_{it} is equal to the year (i.e., 2006 to 2023) if DNSP i is in New Zealand, and zero otherwise; and
- ont_yr_{it} is equal to the year (i.e., 2006 to 2023) if DNSP i is in Ontario, and zero otherwise.

Whereas for the BC95-JTT-HN-GTC model:[41]

- aus_yr_{it} is equal to the year (i.e., 2006 to 2023) minus the average year within the benchmarking period if DNSP i is in Australia, and zero otherwise (i.e., ranging from -8 to +9 if the DNSP is Australian, and zero otherwise);
- nz_yr_{it} is equal to the year (i.e., 2006 to 2023) minus the average year within the benchmarking period if DNSP i is in New Zealand, and zero otherwise (i.e., ranging from -8 to +9 if the DNSP is from New Zealand, and zero otherwise); and
- ont_yr_{it} is equal to the year (i.e., 2006 to 2023) minus the average year within the benchmarking period if DNSP i is Ontarian, and zero otherwise (i.e., ranging from -8 to +9 if the DNSP is Ontarian, and zero otherwise).

And for the BC95-JTT-HN-hY model:

- aus_yr_{it} is equal to the year (i.e., 2006 to 2023) minus the average year if DNSP i is in Australia, and -2014 otherwise (i.e. ranging from -8 to +9 if the DNSP is Australian, and -2014 otherwise);
- nz_yr_{it} is equal to the year (i.e., 2006 to 2023) minus the average year if DNSP i is in from New Zealand, and -2014 otherwise (i.e., ranging from -8 to +9 if the DNSP is from New Zealand, and -2014 otherwise); and
- ont_yr_{it} is equal to the year (i.e., 2006 to 2023) minus the average year if DNSP i is Ontarian, and -2014 otherwise (i.e., ranging from -8 to +9 if the DNSP is Ontarian, and -2014 otherwise).

That is, for reasons that are not clear to us, Quantonomics has defined the efficiency time trend terms differently in different versions of the BC95 model, whereas these time variables ought to be defined consistently across models.

A second, more serious problem is that for some of the BC95 models investigated by Quantonomics—including the BC95 model that Quantonomics identifies as one of the four best-performing models—the efficiency time trend is defined in a way that effectively makes expected inefficiency time-invariant over the benchmarking period.

For example, in the BC95-JTT-HN model (the BC95 model that Quantonomics identified as one of the four best-performing models overall), the pre-truncation mean of the inefficiency term, μ_{it}, will be calculated as $\beta_1 \times 2006$ for Australian observations in 2006, $\beta_1 \times 2007$ for Australian observations in 2007… $\beta_1 \times 2023$ for observations in 2023. All DNSPs are assumed to have $\mu_{it} = 0$ (a half-normal distribution of inefficiency) in year 0, which is over two millennia prior to the start of the benchmarking sample.

This implies that the ratio of μ_{it} in 2023 and 2006 is:

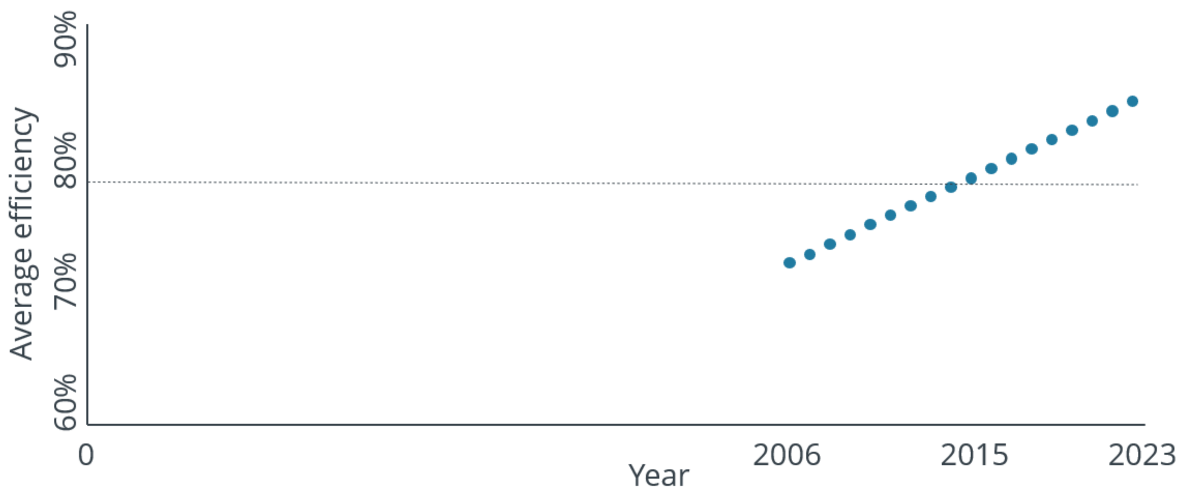
$$\frac{\mu_{i,2023}}{\mu_{i,2006}} = \frac{\beta_1 \times 2023}{\beta_1 \times 2006} \approx 1.$$

[41] A similar approach was applied for the BC95-JTT-HN-GTC-hY, BC95-AJTT-HN-GTC and BC95-AJTT-HN-GTC-hY models.

That is there is almost no difference in expected inefficiency between the end of the benchmarking sample (i.e., 2023) and the start of the benchmarking sample (i.e., 2006).

The problem with the way Quantonomics defined the efficiency time trend in the BC95-JTT-HN model is illustrated in Figure 3 below. The blue dots can be considered (for illustrative purposes) the true efficiency of DNSPs, rising considerably over time from 2006 to 2023. When attempting to fit the data, the Quantonomics specification is essentially constrained to finding the best slope, with the intercept (year 0) set to 77%.[42] Because $t = 0$ is defined to be more than 2,000 years before the start of the benchmarking period, the slope of the line of best fit is essentially flat over that period. Due to the restrictive way in which the efficiency time trend is defined by Quantonomics, the BC95 model that Quantonomics identifies as one of the top-performing models has no hope of estimating the trend in inefficiency accurately (i.e., the line of best fit is almost horizontal, but the true efficiency is trending upwards over the benchmarking period steeply).[43, 44]

Figure 3: Illustrative example – impact of Quantonomics time trend variable on estimated efficiency time trend



Source: Frontier Economics.

To confirm this, we estimated a model that was otherwise equivalent to the Quantonomics' BC95-JTT-HN model, except that instead of using time variables (aus_yr_{it}, nz_yr_{it}, ont_yr_{it}) for the inefficiency component, we used jurisdiction dummies. That is, the expected inefficiency did not vary over time but did vary across jurisdictions.

The resulting estimated cost function is essentially identical to that estimated by Quantonomics, as were the estimated efficiencies – the differences between Quantonomics' estimated efficiencies and

42 This is the expected unconditional efficiency, based on a half-normal distribution (as $\mu = 0$ in year 0 under the Quantonomics specification) and $\sigma_u = 0.360$ as per Quantonomics estimate for the translog model in Table 4.3.1. We note that in theory the intercept can vary to some degree to changes in σ_u, though not to the extent that would allow a line of best fit that adequately matched the true efficiencies as in Figure 3.

43 As noted above, the efficiency time trend in the BC95-AJTT-HN and BC95-AJTT-HN-hY models is defined in the same way as in the BC95-JTT-HN model, so suffer from the same problem that expected inefficiency in those models are essentially time invariant over the benchmarking period.

44 Although the efficiency time trend is defined differently in the BC95-JTT-HN-hY model, it too is essentially a model in which expected inefficiency is time-invariant.

ours were less than 0.02% across all Australian observations. That is, both models yield essentially equivalent results in terms of the cost function and efficiencies. Yet the model we estimated did not allow for a time trend in efficiency – in practice neither did the Quantonomics model.

The comparison of cost function estimates, between Quantonomics' estimates of the BC95-JTT-HN-TL model and a time-invariant model with jurisdiction dummies for the average efficiency, are provided in Table 2 below.

Table 2: Comparison of estimates of frontier cost function

	Quantonomics	Frontier Economics – jurisdiction dummies
ly1	0.448	0.448
ly2	0.103	0.103
ly3	0.415	0.415
ly11	0.334	0.334
ly12	0.026	0.026
ly13	-0.538	-0.538
ly22	-0.005	-0.005
ly23	0.016	0.016
ly33	0.680	0.680
lz1	-0.178	-0.178
yr	0.011	0.011
jur2	0.070	0.070
jur3	0.298	0.298
-cons	-12.368	-12.331

Source: Frontier Economics analysis, Quantonomics analysis

On this basis we conclude that the following models were fundamentally mis-specified, sharing the same treatment of the time trend used in efficiency:

- BC95-JTT-HN;
- BC95-AJTT-HN; and
- BC95-AJTT-HN-hY.

For similar reasons, the BC95-JTT-HN-hY model is also considered mis-specified.

While the four models that use time trends variables that range from -8 to 9 are not mis-specified for the same reason,[45] we do have a concern with the AJTT models within that set.[46] In these specifications, the expected inefficiency is allowed to vary over time, with different impacts of time across Australian DNSPs. Yet the effect of time is constructed in such a way that DNSPs must have the same half-normal inefficiency distribution in the middle of the sample period (i.e., all DNSPs must have $\mu_{it} = 0$ in 2014).

As a consequence, firms had very low efficiency in the beginning of the sample period ($\mu_{it} > 0$), but increased over time, would be evaluated to have higher expected efficiency at the end of the period ($\mu_{it} < 0$), relative to firms that maintained high efficiency throughout the period ($\mu_{it} = 0$ is the only value that can satisfy the condition of constant efficiency over time). This is in contrast to observed efficiency trends, either via the efficiency estimates in Figure 4.2.2 of Quantonomics' Phase 2 report or the Opex MPFP results. The observed pattern is one of convergence in efficiency scores over time, rather than DNSPs with low efficiency scores at the start of the period becoming materially more efficient at the end of the period than those DNSPs that started out with high efficiency scores.

Proposed Frontier Economics solution

The expression for the pre-truncation mean of the inefficiency distribution in the BC95 model can be written:[47]

$$\mu_{it} = \beta_1 aus\_yr_{it} + \beta_2 nz\_yr_{it} + \beta_3 ont\_yr_{it}$$

Our recommended approach is to define the time trend for efficiency in a consistent manner across all the BC95 models. Specifically, we propose an approach whereby:

- aus_yr_{it} is equal to the observation year minus the final year of the benchmarking period if the observation is from an Australian DNSP, and zero otherwise.
- And similarly for nz_yr_{it} and ont_yr_{it}.[48]

For example, if the benchmarking period was 2006 to 2023, then the time trend term for an Australian DNSP would be:

- -17 (i.e., 2006 – 2023) for the first year in the benchmarking period;
- -16 (i.e., 2007 – 2023) for the first year in the benchmarking period;
- …
- 0 (i.e., 2023 – 2023) for the final year in the benchmarking period.

The efficiency time trend term would be defined analogously for the New Zealand and Ontarian DNSPs.

---

45 These models are: BC95-JTT-HN-GTC, BC95-JTT-HN-GTC-hY, BC95-AJTT-HN-GTC and BC95-AJTT-HN-GTC-hY.

46 These models are: BC95-AJTT-HN-GTC and BC95-AJTT-HN-GTC-hY.

47 An intercept/constant term is not included, in line with Quantonomics' implementation of the half-normal assumption in the BC95 model.

48 Note that we adopt the yr variable used by Quantonomics, which adjusts the year for differences in the final month of the reported year. We also subtract 2022 for Ontario DNSPs for similar reasons.

This specification allows a meaningful impact of time on the expected inefficiency of DNSPs (because the change in the time trend term from year to year is meaningfully large), while also applying a half-normal distribution for all Australian DNSPs in the final year of the benchmarking period. That is, in the example above, since $aus\_yr_{i2023} = 0$, therefore $\mu_{i2023} = 0$.

This has two benefits:

- The approach allows for convergence of inefficiency distributions across DNSPs at the end of the sample period, rather than in the middle of the period (i.e., convergence followed by divergence), which is more consistent with the observed pattern of changes in the Opex MPFP results; and
- When evaluating the efficiency of DNSPs in the final year, the estimation is less sensitive to the trajectory of past performance of the DNSP i.e., all DNSPs are compared on an equal basis in the final year of the benchmarking period.

The second point is in our opinion a prerequisite for estimates of efficiency in the final year to be used to compare DNSPs, either for the purpose of evaluating base year opex or simply for comparing/ranking DNSPs. If not, resulting estimates of efficiency of DNSPs that have experienced considerable catch-up efficiency gains may be biased downwards.

We note that each of the Kumb90 models estimated by Quantonomics use a time trend for efficiency adjustments as in the BC95-JTT-HN-GTC model: ranging from -8 to +9 if the DNSP is located in the relevant jurisdiction, and zero otherwise. In the context of the model (including that the efficiency adjustment equation retains an intercept term) we do not object to the formulation and adopt that for our evaluation of models in section 5.

3.2 Kumb90 model efficiency time trend is overly restrictive

Explanation of the issue

In the Kumb90 models considered by Quantonomics, each DNSP draws a time invariant inefficiency term u_i once. The actual inefficiency in each year, u_{it}, applies a deterministic adjustment factor to the DNSP's u_i. In the first two specifications considered by Quantonomics, Kumb90-JTT-HN and Kumb90-JTT-HN-GTC, the adjustment factor is the same for all Australian DNSPs in any given year. This has several implications:

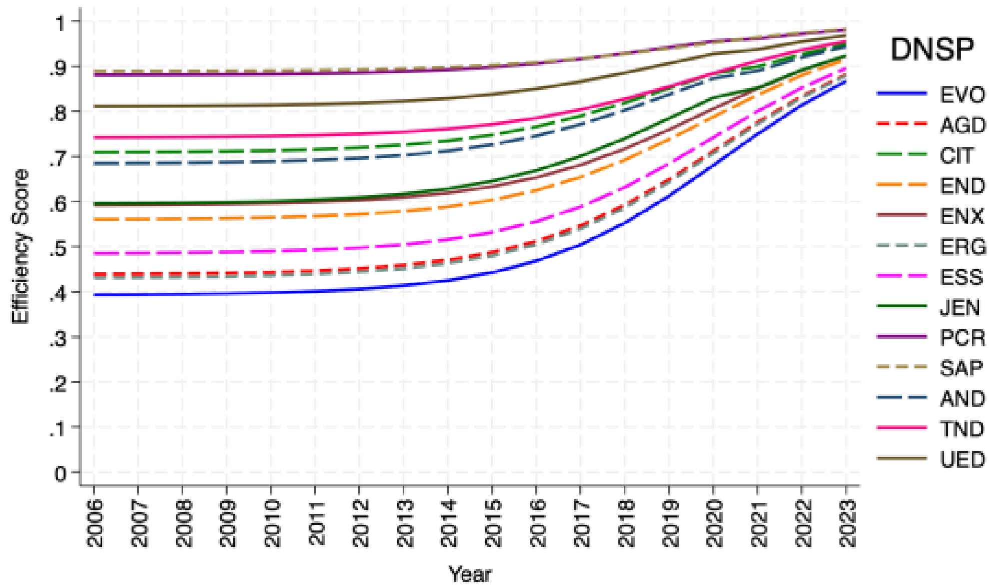
- Efficiency scores of DNSPs will converge if efficiencies tend to change monotonically over the sample period;
- Efficiency scores will converge to 100% if efficiencies tend to increase (converging to some level <100% if decreasing over the sample period); and
- The efficiency score rank of a DNSP will never increase or decrease.[49]

The first two points above, regarding the convergence of efficiency scores, can be seen in Figures 4.4.2 (reproduced below, for illustrative purposes in **Figure 4**) through 4.4.4 of the Quantonomics Phase 2 report. There is considerable convergence, with around 80% of variation in efficiency scores eliminated between 2006 and 2023. This seems rather extreme and inconsistent with the observed

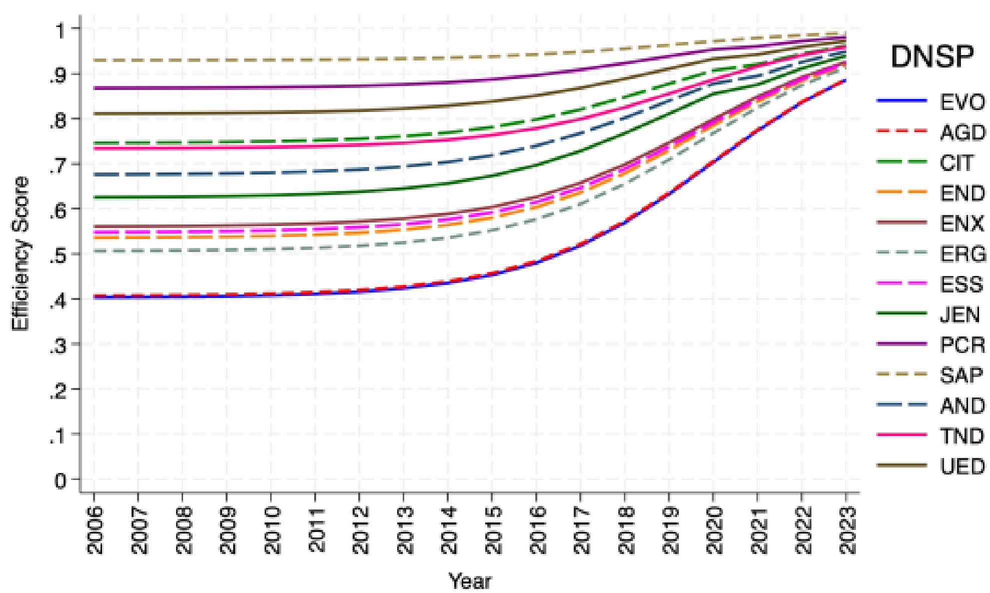
49 We note that Jemena appears to have switched with Energex twice – this was simply due to reporting in calendar vs fiscal year, and the switch in 2021 rather than an actual change in ranking.

trends in the trends in the Opex MPFP indices (see **Figure 5** below). The convergence also appears to accelerate in later years, although this is a consequence of the specification of the adjustment factor rather than reflecting any pattern in the actual speed of adjustment over time.

Figure 4: Efficiency trends by DNSP under the Kumb90-JTT-HN model estimated by Quantonomics



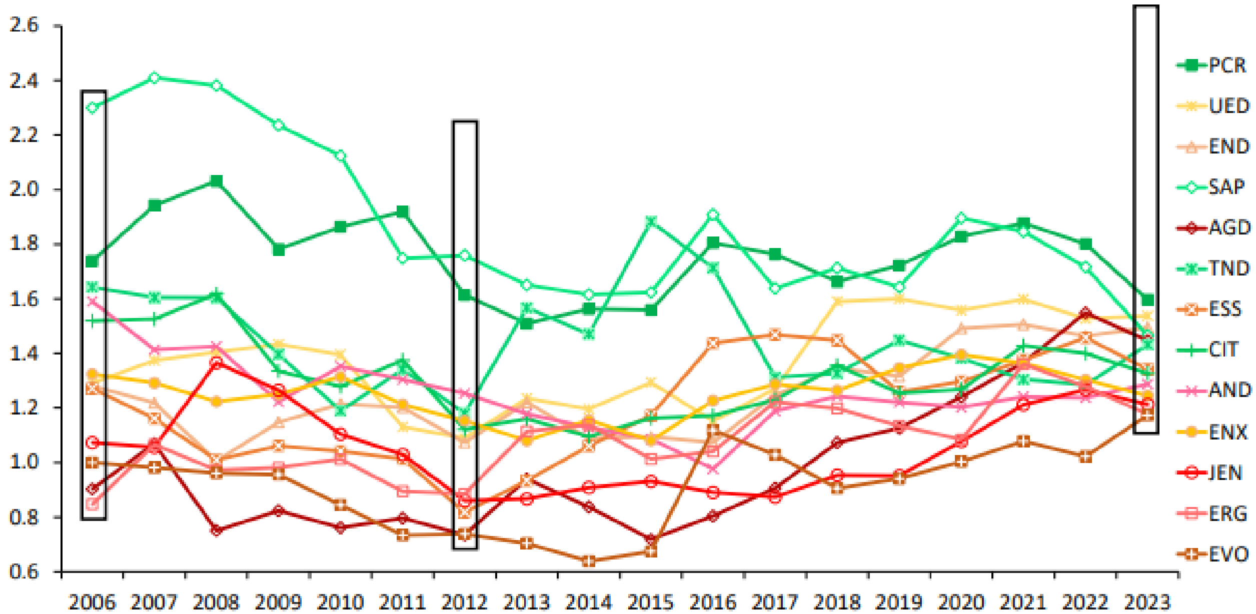
SFACD Long Period



SFATLG Long Period

Source: Figure 4.4.2 in the Quantonomics Phase 2 report.

Figure 5: Opex MPFP indices by DNSP



Source: AER 2024 Annual Benchmarking Report for DNSPs, Figure 12.

The convergence towards 100% efficiency is also problematic – when the sample period becomes longer in the future, the model estimated on the longer sample period may fit poorly if actual efficiency does not increase as considerably as suggested by the current results.

The final implication, that the rank will never increase or decrease, is problematic from the perspective of application of the efficiency estimates. The impact of a DNSP's opex on estimated efficiency in a given year is minimal and is essentially limited to the impact on u_i—i.e., its average efficiency over the sample period. Thus, if a firm substantially increases (or decreases) its efficiency in the final year of the sample, the resulting estimate of efficiency in the final year will be slightly higher (lower) than it would otherwise, and all previous years will similarly have higher (lower) efficiency. That is, the estimates of efficiency for individual years are uninformative of actual efficiency in those years.

The low level of responsiveness of individual year efficiency estimates to efficiency improvements is problematic regarding incentives to reduce opex. The responsiveness to actual efficiency improvements is similar to that of the current regime. Firms that materially increase their efficiency relative to peers are unlikely to see material improvements in efficiency estimates as a result, and will see little prospects for improving their ranking when examining efficiency in the final year of the sample.

Proposed Frontier Economics solution

Our primary concern with the Kumb90 models estimated by Quantonomics is the overly restrictive pattern of efficiency over time. We propose a less parameterised specification of the efficiency adjustment term that is more flexible in how it changes over time. In section 4.3 we specify a version we term General Efficiency Change (GEC), as an analogue to the GTC approach used for the cost function, but applied to the Australian DNSPs in the efficiency adjustment equation.

As per the GTC approach proposed by Quantonomics, we divide the sample years into groups of three years, defined in the same way as the GTC variables, with each group having the same efficiency adjustment factor for Australian DNSPs. We retain the Quantonomics approach for the New Zealand and Ontarian DNSPs for simplicity.

This approach does not restrict the shape of adjustment over time as per the current approach, does not guarantee convergence in the same way, and does not require efficiency scores to converge to 100%.

3.3 LSE models assume a constant rate of efficiency catch up

The traditional justification that has been given for use of LSE models relates to the closed-form nature of the least squares estimators, avoiding the convergence issues they encountered with the SFA models when attempting to incorporate time-varying efficiency. Quantonomics states that:

> *unlike the SFA models, the LSE time-varying models proved to be more robust in estimation and can reliably include DNSP-specific efficiency trends.*[50]

Quantonomics argues that the ease of calculation of the LSE models relates to their relative simplicity:

> *In contrast, the LSE time-varying models proved to be more robust and consistent across specifications and time periods. This reflects their relative simplicity compared with the SFA models, which are more complex as the stochastic term comprises two components, white noise and inefficiency, with the latter identified through the composed-error formulation. The LSE models have a simple white noise disturbance, permitting a more straight-forward computation. Hence, some of the LSE models that include DNSP-specific efficiency trends are reasonably reliable and can be considered candidate models.*[51]

The LSE models that Quantonomics investigated assume a constant rate of efficiency change over the whole benchmarking period.[52] This allows no possibility for DNSPs to catch up more or less quickly over time,[53] which is a strong assumption and does not align with the highly nonlinear trajectories of the Opex MPFP indices. By contrast, the BC95 and Kumb90 models can accommodate (to varying extents) nonlinear efficiency trends among Australian DNSPs and jurisdictions.

However, the LSE models fit the data poorly, as indicated by the residual plots for those models:

- LSE residuals exhibit much higher variance than those from the SFA models, which indicates poorer fit; and

50 Quantonomics Phase 2 report, p. v.

51 Quantonomics Phase 2 report, p. 79.

52 Though allowing for differences in efficiency time trends across Australian DNSPs.

53 Holding the frontier firm constant. The rate of change in efficiency changes when a new DNSP becomes the reference DNSP, as seen in Figure 4.5.2 of the Phase 2 report.

- LSE residuals vary systematically across time periods, suggesting that the linear efficiency trend assumption is violated (see Appendix A). This can result in severe mis-specification issues: the time-varying residuals may become correlated with the observed covariates, resulting in biased cost-function coefficients.[54]

In view of this, the relative simplicity involved in estimating the LSE models should not be viewed as a clear advantage over the SFA models. The LSE models may be simpler to estimate, but that comes at the cost of poorer model fit. Since we show (in the next section) that there are no instances of the BC95 and Kumb90 models failing to converge, when appropriate starting values are used, there is no benefit in trading off model fit for simplicity in the estimation process.

3.4 Model convergence

Explanation of the issue

Quantonomics notes in the Phase 2 report that, of seven of the eight BC95 models it considered failed to converge. As a consequence, Quantonomics considered these models as unsuccessful, and concluded that "the BC95 model becomes more difficult to estimate as the models become more complex, and with the ABR data, estimation was only feasible in the simplest of the model specifications tested."[55]

In our evaluation of the modelling, we found that Quantonomics did not set reasonable starting values for the estimation of the SFA models, which is what gave rise to the apparent convergence issues encountered by Quantonomics. In other words, the models were unable to arrive at sensible estimates as they did not start at sensible estimates.

In previous estimation of SFA models, Quantonomics has adopted starting values for the distribution of inefficiency and noise, in particular the variance of the two terms.[56]

In the Stata .do file Quantonomics uses to estimate the BC95 models, it applies the following procedure:

- Estimate an LSE version of the cost function;
- Take the resulting cost function estimates, along with the log of error variance, as starting points for a simple SFA model (a value of 0.1 is used for the relative size of inefficiency and idiosyncratic error variance);
- Estimate a simple half-normal SFA model (no inefficiency effects);
- Take the resulting cost function, along with starting coefficients of 0.0001 for the impact of aus_yr_{it}, nz_yr_{it} and ont_yr_{it} on inefficiency; and
- Estimate the full BC95 model.

54 Note that LSE models rely heavily on within-firm, time-varying covariates for identification of cost-function parameters.

55 Quantonomics Phase 2 report, p. 50.

56 Note that the specification used in the *xtfrontier* package uses the log of the variance of the combined term, as well as a parameter showing each component's contribution to the total variance.

We see two key issues with that approach:

- In the estimation of the full BC95 model, no starting values are provided for the variance of the inefficiency and idiosyncratic error terms even though useful starting values for these are provided in the intermediate SFA estimation. This is in contrast to Quantonomics' previous practice of using starting values for these components in SFA estimation; and
- When estimating the translog models, which tend to be more challenging to estimate, Quantonomics does not appear to consider simply starting with the estimates from the relevant Cobb-Douglas specification (both the cost function and SFA parameter estimates) and using starting values of zero for the translog terms. This allows the more challenging components (the SFA parameters) to be estimated on an easier model, and starts the translog estimation at starting points that provide relatively good fit.

We also see little benefit in using starting points that imply that inefficiency rises over time (i.e., a positive coefficient) and suggest setting the impacts of the jurisdiction time trends on inefficiency to be zero. This aligns the starting values with the time-invariant efficiency model used to generate starting values of other model components.[57]

Proposed Frontier Economics solution

The approach we propose for obtaining appropriate starting values involves:

- First estimating a simple least squares Cobb-Douglas model;
- Using the estimated cost function, along with the error variance, as starting values in estimating a time-invariant SFA model, with a half-normal inefficiency distribution;
- Using the estimated cost function, along with the inefficiency/idiosyncratic error variances, as starting values in estimating a time-varying Cobb-Douglas SFA model, with the inefficiency μ term allowed to vary with time (though setting starting values to have no trend i.e. $\mu_{it} = 0$ for all DNSPs/years); and
- Using the estimated cost function, along with the inefficiency/idiosyncratic error variances and efficiency time trends, as starting values in estimating a time-varying translog SFA model.

The details of this approach are explained in Appendix A. When we apply this approach, we have no difficulty in obtaining convergence for all models estimated, including the eight BC95 models considered by Quantonomics.

Accordingly, we disagree with the suggestion that estimation is only feasible for the simplest of the BC95 model specifications tested. As a consequence, we consider that more complicated models should be examined, where appropriate. For example, we consider relaxing the linear component of the efficiency trend in a similar manner to the GEC modification to the Kumb90 models, discussed above.

57 The standard time-invariant efficiency model can be considered as the time-varying BC95 model but restricting the time trend in efficiency to zero.

3.5 Calculation of efficiency estimates under some models

Explanation of the issue

Quantonomics observes that some SFA models obtain convergence yet do not result in valid efficiency estimates for some observations. This issue was previously highlighted in Frontier Economics modelling provided to the AER.[58]

In such cases, the estimation of the model has been successful, in that the parameters of the cost function and inefficiency/noise distribution terms have been successfully estimated. The efficiency estimates are then sought using post-estimation procedures. Given the cost function estimates one can evaluate the distance between actual costs and the efficient frontier for each observation and use this to form an expectation of efficiency.[59]

We noted that in some circumstances, the distribution of inefficiency may be so extreme that the standard Stata command encounters a division-by-zero error in the calculation of the expected efficiency score. This happens when, for example, the pre-truncation mean of inefficiency μ_{it} is highly negative, such that the inefficiency distribution is clustered near 0. This division-by-zero error results in no efficiency estimate being generated by the standard Stata command.

Proposed Frontier Economics solution

We have reviewed the relevant Stata .ado file of the *sfpanel* package, as used by Quantonomics estimate efficiency. We have replicated the calculation, except using the *lnnormal* command to avoid the division-by-zero error. By using the *lnnormal* command we are able to avoid issue of probabilities being rounded down to zero as a result of the format in which Stata stores numbers. The calculations are provided in Appendix B. The resulting efficiency estimates are exactly the same as those generated by the standard Stata post-estimation command, except they generate a valid efficiency estimate in all cases—even those instances in which Quantonomics was unable to derive an efficiency estimate using the standard Stata command. Our approach is trivial to implement.

Quantonomics claims that a "problem of being unable to produce efficiency scores for all the Australian DNSPs is sufficient grounds for rejecting" a particular specification. Our method is able to produce valid efficiency scores for all Australian DNSPs; there is no need to reject models on that basis. The required calculations are not an approximation – they generate the expected efficiencies using the same equation as the built-in command and in fact replicate the values.

Finally, we note that the issue is typically encountered in the class of models referred to as AJTT within the BC95 class. In such models. where individual DNSPs have their own efficiency time trend: some highly efficient DNSPs have a time trend that results in extreme distributions of efficiency that result in the division-by-zero error. We have not encountered this issue in Kumb90 models and would not expect to do so because in Kumb90 models all DNSPs draw the time invariant component of efficiency from the same distribution.

[58] See rows 233-254 of "2025-04-14 SF models.do", provided to the AER during the Quantonomics Phase 1 memorandum consultation.

[59] When using the *sfpanel* package in Stata this is accomplished by the command "predict efficiency, bc".

3.6 Consideration of omitted variables

Section 2.2.4 explained that Quantonomics explored a number of potential omitted explanatory variables from the econometric benchmarking models but ultimately concluded that none of the variables investigated should be included in the models.

A potentially important issue that was not investigated by Quantonomics is whether any new output variables should be included to account for the changing role played by DNSPs in the energy system. Historically, DNSPs' primary role was to step down voltage and to transport electricity safely and reliably from transmission systems to consumers, with little active control over power flows or demand. Traditionally, electricity flowed one way from large, centralised generators, through transmission systems, into distribution feeders and, finally, to end-users who were passive recipients of power.

However, over time and with the development of new technologies, the energy system—and the role played by DNSPs—has become more complex and dynamic. Over the last decade in particular, there has been significant growth in consumer energy resources (CER), such as rooftop solar, batteries, electric vehicle chargers and controlled loads. By way of example, **Figure 6** and **Figure 7** show a steep rise over time in cumulative solar PV capacity and small-scale solar installations by State and for the NEM, underlining this trend.

Supporting the delivery of CER services has led DNSPs to undertake new network and non-network investments, which incur ongoing operating and maintenance costs that the current benchmarking models do account for. The facilitation of two-way power flows and delivery of CER services are arguably new DNSP outputs that are not recognised in the existing benchmarking models. However, the opex associated with those services is reflected in the opex that is benchmarked using those models.

In addition, DNSPs are increasingly playing a role in supporting grid stability—particularly as CER has grown—including through load control, fault protection, demand management and investment in community batteries. Grid stability is valuable to consumers because it is a prerequisite for the reliable delivery of power to users. Arguably, the provision of grid stability is also an important output that is not accounted for in the existing benchmarking models, but that require expenditure by DNSPs.

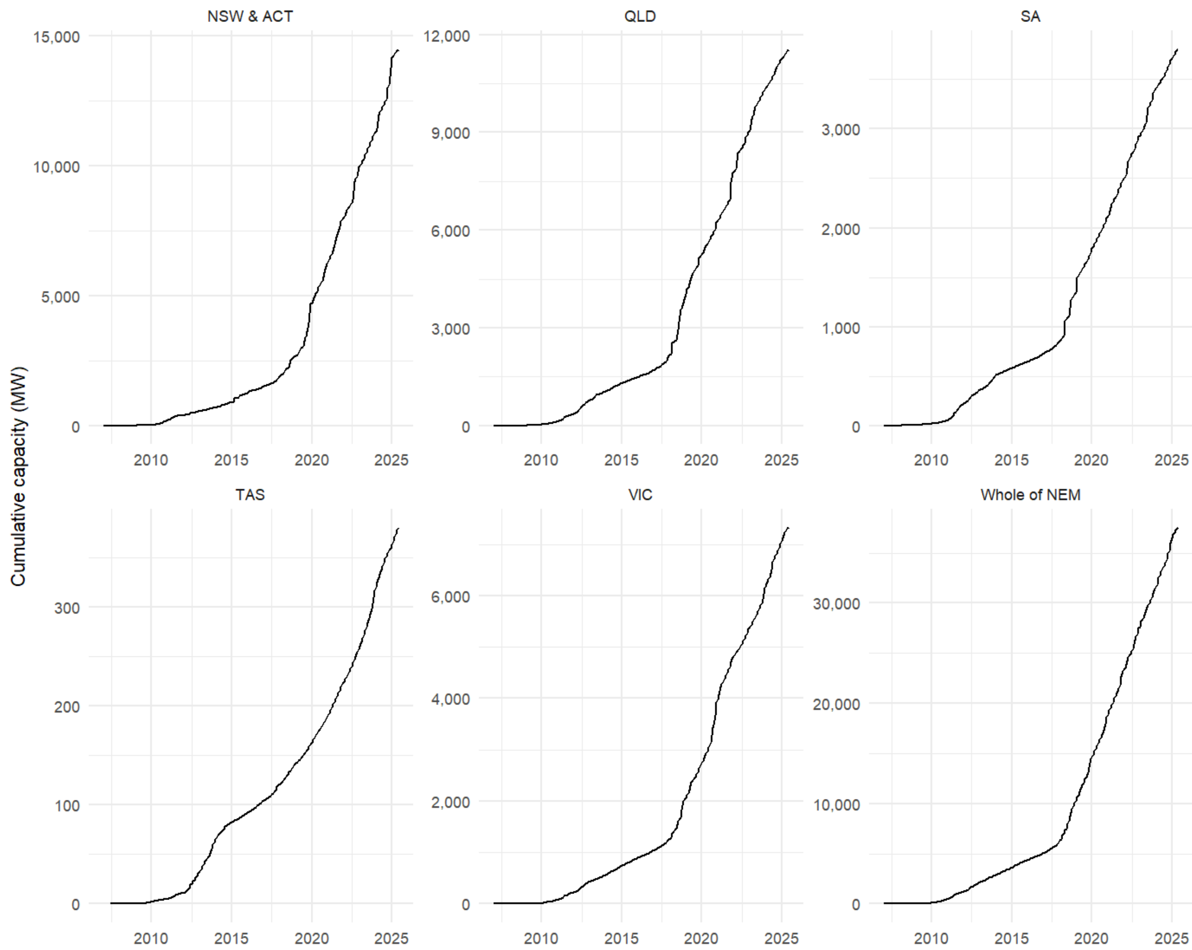
If these are genuinely important outputs that are not accounted for in the models, then the resulting estimates of efficiency for individual DNSPs, and of the cost function used to estimate efficient opex, will be unreliable.

However, there are challenges accounting for these potentially important outputs in the models:

- In order to be incorporated into the models, the outputs must be clearly defined and measurable. However, outputs such as the provision of CER services and supporting grid stability are multi-faceted, meaning that there is unlikely to be one single measure that captures the effect of the output completely.
- Secondly, even if measurable outputs could be defined (e.g., through identification of suitable proxy variables), it may be challenging to obtain consistent measures of those outputs across all DNSPs. The AER could develop a consistent reporting template for Australian DNSPs. However, the benchmarking sample used by the AER is dominated by networks overseas, and there is no guarantee that consistent data can be collected for all the New Zealand and Ontarian DNSPs. By way of example, based on preliminary investigations, we found that there are no consistent data

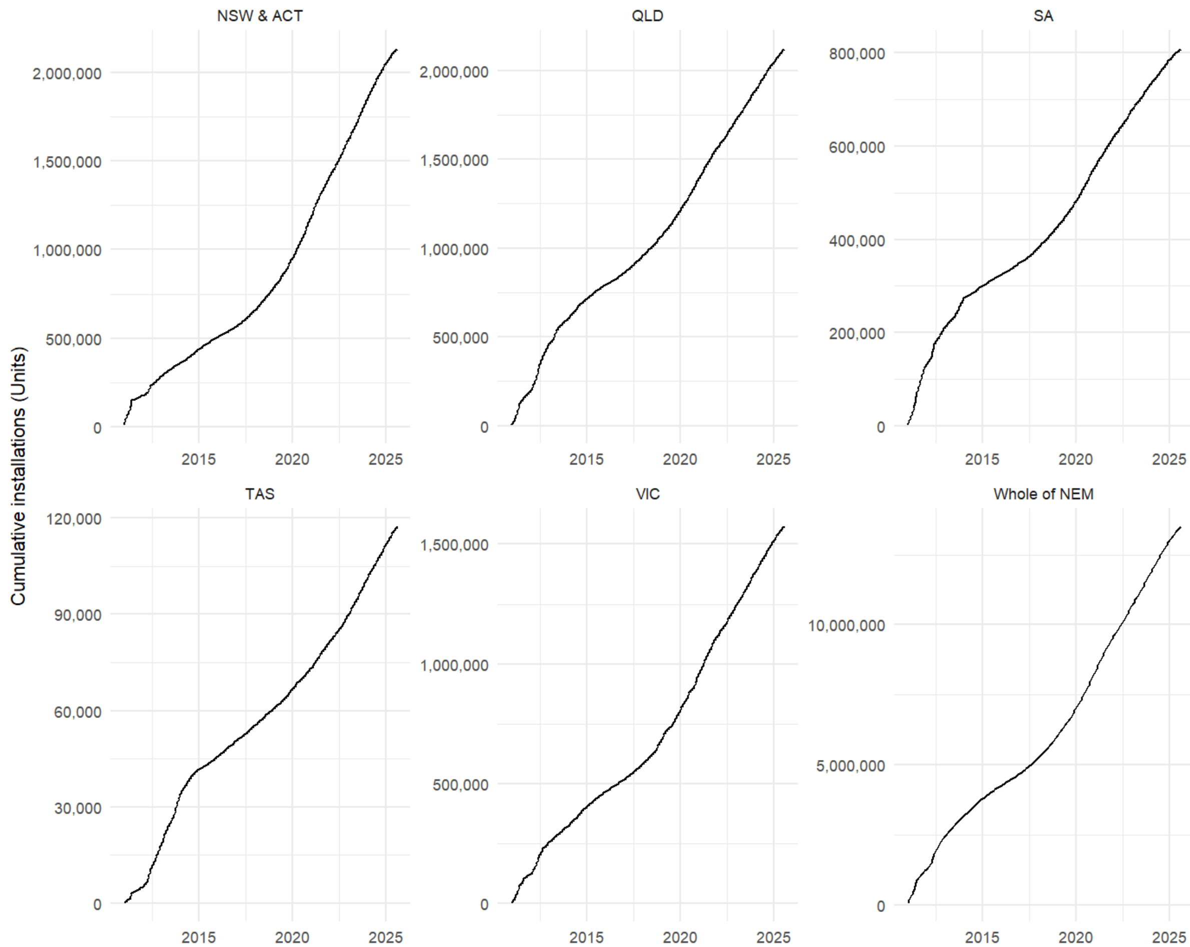
across the three jurisdictions on the volume of solar exports or on investment in network capacity to support solar exports (potential proxy variables for the output of facilitating two-way power flows).

Figure 6: Cumulative capacity of installed solar PV by State and for the NEM



Source: Frontier Economics analysis of data from the Australian Photovoltaic Institute.

Figure 7: Cumulative installations of small-scale solar PV by State and for the NEM



Source: Frontier Economics analysis of data from the Clean Energy Regulator.

Given these challenges, there seems to be little that could be done in the short-term to account for these potentially important omitted variables. However, we recommend that:

- The AER undertake further work over the medium-term (as part of its ongoing development of the benchmarking models) to clearly define the new outputs being provided by DNSPs, and to investigate whether these outputs can be quantified in a reliable and consistent way, such that they can be incorporated into the benchmarking analysis; and
- In the meantime, the AER should recognise that the inability to account for these outputs (which are likely to grow in importance over time) is a further limitation of the benchmarking analysis such that the results of the benchmarking analysis should be treated with caution, and not used deterministically, when making revenue determinations for DNSPs.

4 Models we investigated

4.1 Overview

As set out in section 3, we proposed improved specifications of the BC95, Kumb90 and LSE models (relative to those presented in the Quantonomics Phase 2 report). We also proposed an approach for obtaining appropriate starting values that ensures convergence for all the SFA models estimated (including those that Quantonomics found did not converge), and a method to obtain efficiency scores for all Australian DNSPs under all the models tested.

In this section summarise the time-varying inefficiency models that we investigated. This includes models investigated by Quantonomics, but modifying the specifications to align with the approaches we proposed in section 3.

We also investigated additional specifications not considered by Quantonomics. These models were extensions to, or modifications of, the models considered by Quantonomics—for example to allow for alternative treatments of the time trend for efficiencies under the different models.

In all cases we apply our approach to setting starting values, to overcome the convergence issues encountered by Quantonomics and employ the post-estimation routine we recommend in in section 3 that ensures we are able to derive efficiency scores for all DNSPs under all models.

Note that for the purpose of comparability, we rely on the same benchmarking dataset used in the Quantonomics Phase 2 report: the data employed in the 2024 Annual Benchmarking Report, which includes historical data between 2006 and 2023 (inclusive).

Whilst there are many time-varying inefficiency models that could, in principle, be investigated, we focus on the BC95, Kumb90 and time-varying LSE models considered by Quantonomics, on the basis that these models:

- can be estimated using the existing datasets;
- are relatively minor extensions to the models currently used in benchmarking, but allow extensions to address time-varying efficiency; and
- are established in the literature and are included in the well-documented Stata package *sfpanel*, which makes their implementation relatively straightforward and practical.

We restricted our testing of the SFA models to the translog (rather and Cobb-Douglas) specifications, since the translog models are highly flexible and, therefore, more prone to statistical problems such as monotonicity violations.[60]

We also restrict the analysis to the long benchmarking period (i.e., from 2006 to 2023) because, once again, it is the models estimated over this longer period that have historically been found by the AER to be unreliable (e.g., due to excessive monotonicity violations) and therefore discarded. If the most flexible models perform well over the long benchmarking period, they are more likely to be robust models that the AER could adopt going forward. The longer sample increases the susceptibility of the

[60] In addition, as set out in Section 5.3, we ultimately recommend focussing on the use of translog models, with Cobb-Douglas models serving as a backup option rather than an equally valid source of estimated efficiency/efficient base year opex.

model estimates to changes in efficiency, and so provides a stronger test of the capacity of the model to handle such changes now and in the future. This is consistent with the approach of Quantonomics: "The discussion focuses primarily on the long-period specifications of the time-varying inefficiency models, as these account directly for changes in inefficiency and therefore reduce the reliance on the short-period results."[61]

4.2 BC95 models

Models we investigated

We consider the four variants considered by Quantonomics:

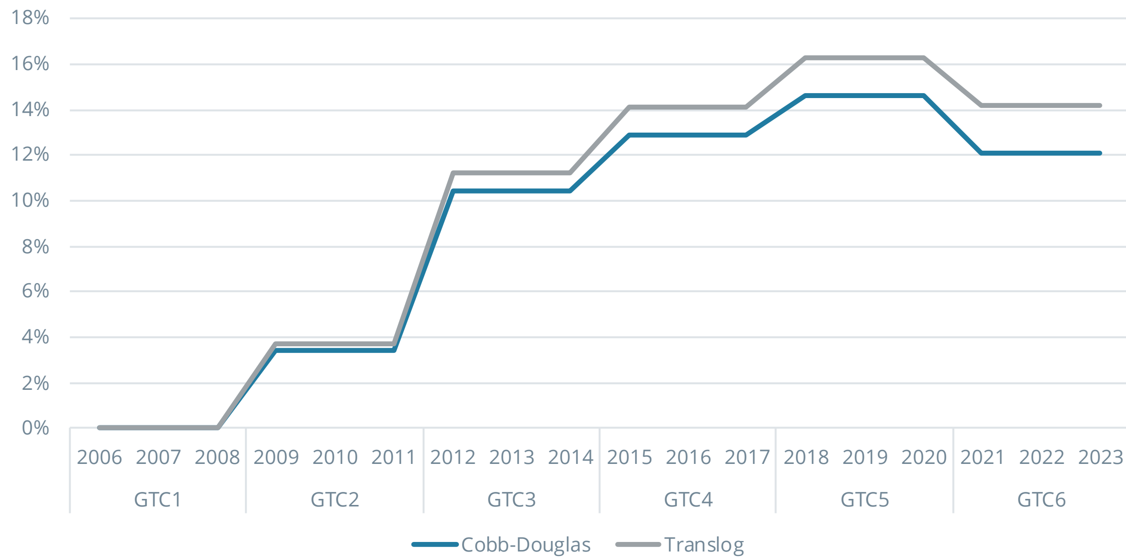
- BC95-JTT-HN;
- BC95-JTT-HN-GTC;
- BC95-AJTT-HN; and
- BC95-AJTT-HN-GTC.

As noted in Section 2.2.1, Quantonomics estimates 'half-normal' variants of the BC95 models. By omitting the constant term of the mean of the non-truncated inefficiency distribution Quantonomics seeks to simplify the estimation process, assisting in convergence/identification of the models. Whilst this is more restrictive that the truncated-normal assumption underpinning the existing SFA models, we think this is a reasonable restriction to impose, given the improved ability to estimate the models numerically.

We consider that the GTC extension proposed by Quantonomics, which allows a more flexible impact of time on the cost function, is a sensible extension to the linear time trend in the most basic versions of the three models. In particular, some estimates of the GTC time trends suggest that the impact of time on cost is not linear and instead there was a period of substantial real cost growth followed by lower growth or even cost reductions. Such a pattern is evident from the estimated time trends for the LSE-ADTT–GTC model illustrated in **Figure 8**. The Figure suggests that costs trended up sharply between the 2009-11 and 2012-14 periods. This was followed by a relatively slow increase in costs and, eventually, a reduction in costs between the 2018-20 and 2021-23 periods.

61 Quantonomics, Phase 2 report, p.38.

Figure 8: Time trend estimates over different time periods in the LSE-ADTT–GTC model



Source: Frontier Economics analysis.

Efficiency time trend

The JTT specification of the efficiency time trend, while relatively simple and linear in latent index,[62] appears to be useful as a starting point for modelling time varying efficiency within the BC95 framework. The AJTT specification, an extension allowing the efficiency time trend to vary across Australian DNSPs (while retaining common efficiency time trends for New Zealand and Ontario DNSPs) is a relatively minor extension that is flexible in allowing for differences in how inefficiency changes over time for Australian DNSPs. Accordingly, we consider both options.

We do observe that efficiency growth may not have occurred in a manner entirely consistent with the linear trends of the JTT and AJTT specifications. In particular, there seems to be relatively little efficiency improvement among Australian DNSPs between 2006 and 2015—the period preceding the introduction of formal economic benchmarking within the revenue reset process for DNSPs.

Subsequently, substantial improvements were observed for a number of DNSPs, consistent with introduction of the AER's current benchmarking regime. To account for this pattern, we adopt a more flexible General Efficiency Change (GEC) approach, using a different efficiency for each three-year period in the sample, analogous to the GTC approach of the cost function. However, for simplicity and parsimony, we apply the GEC approach to the Australian DNSPs only; we retain the simple linear efficiency time trend approach for New Zealand and Ontarian DNSPs.

Note that this approach requires the estimation of four additional parameters: five parameters are estimated for the six GEC variables (excluding one category) instead of the single linear term for Australia.

62 Linear in how time impacts the mean on the non-truncated efficiency distribution, rather than linear in terms of the impact of time on efficiency.

When estimating this variant, we adopt only the GTC specification of the cost function time trend. The fifth specification is therefore referred to as the BC95-AGEC-HN-GTC model, which has the following functional form:

$$u_{it} \sim N^+(\mu_{it}, \sigma_u^2)$$

With the mean of the non-truncated distribution calculated as:

$$\mu_{it} = \delta_1 ausGEC1_{it} + \delta_2 ausGEC2_{it} + \delta_3 ausGEC3_{it} + \delta_4 ausGEC4_{it} + \delta_5 ausGEC5_{it} + \delta_6 nz\_yr_{it} + \delta_7 ont\_yr_{it}$$

Where $ausGEC1_{it}$ is equal to 1 if the DNSP is Australian and the year is in the GEC1 period (2006 through 2008), and similarly for $ausGEC2_{it}$ through $ausGEC5_{it}$.

We note that Quantonomics also considered variants of these four models allowing heteroscedasticity – the variance of the idiosyncratic error term was allowed to vary based on the 'scale' of the DNSP. We do not consider these models in our analysis, noting that Quantonomics did not find this extension to be promising.[63]

As noted in Section 3.1, we take the approach that the value of the time variable used for the efficiency component to be zero in the final year. For the BC95-AGEC-HN-GTC specification, the analogous approach is to omit the GEC6 variable (i.e., years 2021-2023) from the efficiency equation, so that Australian observations during the final year have a half-normal distribution of inefficiency, as in the JTT and AJTT specifications.[64]

4.3 Kumb90 models

Models we investigated

We consider the two JTT variants considered by Quantonomics:

- Kumb90-JTT-HN; and
- Kumb90-JTT-HN-GTC.

We also consider the AJTT variants considered by Quantonomics:

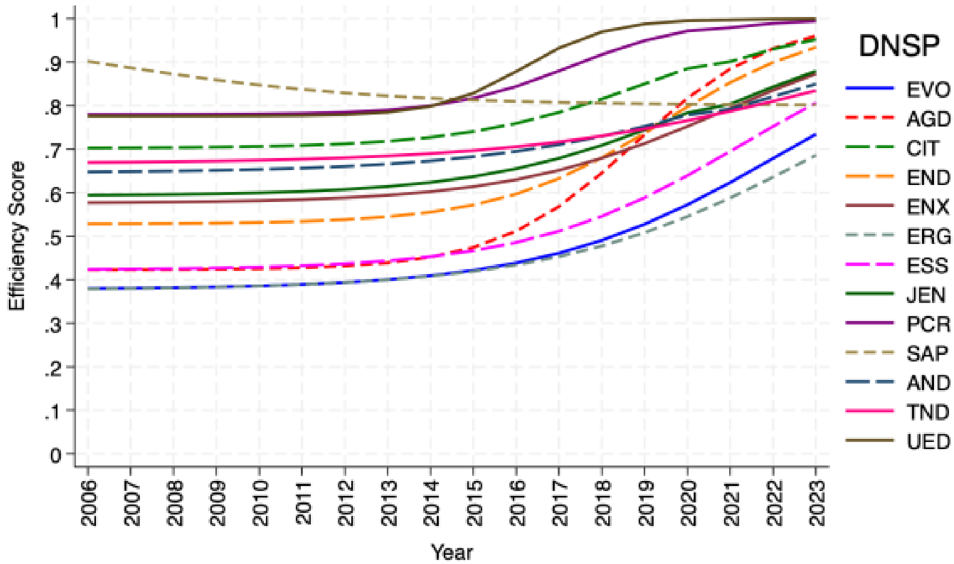
- Kumb90-AJTT-HN; and
- Kumb90-AJTT-HN-GTC

The intent of these models is to allow different changes in the efficiency adjustment factor over time for different Australian DNSPs. Yet the restrictiveness of the specification is apparent when examining Figure 4.4.5 in the Quantonomics Phase 2 report (reproduced below in **Figure 9**), with implausible efficiency trends for Essential Energy over the long benchmarking period, among other DNSPs. While the AJTT version is less restrictive than the JTT version, the formulation of the Kumb90 efficiency adjustment (as applied by Quantonomics) appears too restrictive to yield intuitive or plausible trends in efficiency over time.

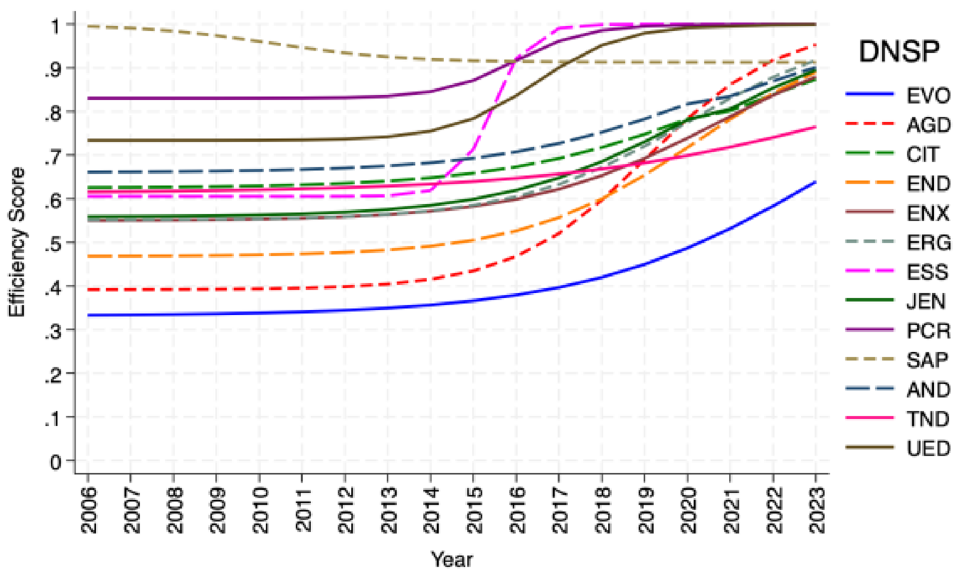
63 Quantonomics, Phase 2 report, p. 50.

64 This would of course be changed as additional years, and therefore GEC variables, are added to the model.

Figure 9: Efficiency trends by DNSP under the Kumb90-AJTT-HN model estimated by Quantonomics



SFACD Long Period



SFATLG Long Period

Source: Figure 4.4.5 in the Quantonomics Phase 2 report.

In addition, we consider GEC versions of the models for the efficiency time trend for Australian DNSPs, while retaining the linear time trend for DNSPs in New Zealand and Ontario. The specification of the efficiency time trend under this approach is:

$$g_{it} = [1 + \exp(\delta_0 + GEC_{it}'\beta_{GEC} + \beta_2 Tnz_{it} + \beta_3 Tont_{it})]^{-1}$$

We also consider a specification that includes jurisdiction dummies for New Zealand and Ontario in the efficiency adjustment term:

$$g_{it} = [1 + \exp(\delta_0 + GEC'_{it}\beta_{GEC} + \beta_2 Tnz_{it} + \beta_3 Tont_{it} + \beta_4 NZ_i + \beta_5 Ont_i)]^{-1}$$

This addition allows further differences in the pattern of efficiency over time across jurisdictions.

Note that each of these approaches are still estimated as half-normal SFA models. The DNSPs draw from the same half-normal inefficiency (pre-adjustment) distribution.

For each of these GEC approaches we consider both linear cost trend and GTC versions. We denote the four additional models as:

- Kumb90-AGEC-HN;
- Kumb90-AGEC-HN-GTC;
- Kumb90-AGECJUR-HN; and
- Kumb90-AGECJUR-HN-GTC,

where 'AGEC' denotes the use of the generalised efficiency trend variables for Australian DNSPs (other DNSPs retaining the simple linear trend) and JUR denotes the inclusion of jurisdiction dummies in the efficiency adjustment equation.

Efficiency time trend

When using linear time trends for the efficiency adjustment, we use the time trend (variable) that is set to zero in the middle of the sample period, consistent with the approach of Quantonomics for the estimation of JTT Kumb90 models.

We apply a similar approach in defining the DNSP specific time trends when estimating AJTT models, though we note that Quantonomics estimated those models by using a DNSP-specific time trend that equals zero at the start of the sample period.

In the Kumb90 class of models there is limited benefit of using a time trend equal to zero in the final period, in contrast to the BC95 models. This is because all DNSPs draw a single efficiency term from the same distribution, to which an identical adjustment factor is applied each year. Instead, we find that models tend to perform better when using the centred time trend.

When using the GEC models, we omit the final GEC term from the estimation, so that the final period/s serve as reference periods. We note that this is irrelevant when using the specification with jurisdiction dummies in the efficiency adjustment equation – the additional dummy variables lead to an equivalent model fit regardless of which term is omitted.

4.4 LSE models

Models we investigated

We consider the two ADTT variants of the time-varying LSE models investigated by Quantonomics:

- LSE-ADTT; and
- LSE-ADTT-GTC.

We also consider the AJTT variants investigated by Quantonomics, noting that Quantonomics includes the New Zealand (rather than Ontarian) time trend in these models:

- LSE-AJTT-NZ; and
- LSE-AJTT-NZ-GTC

We note that in adding a New Zealand time trend to the AJTT variants of the time-varying LSE models, the base time trend is that of the Ontarian DNSPs.[65] It is not immediately clear why a specification with a New Zealand time trend should be favoured over a model with an Ontarian time trend. Accordingly, we investigated an alternative set of AJTT models that included the Ontarian rather than New Zealand time trend:

- LSE-AJTT-Ont; and
- LSE-AJTT-Ont-GTC

If the Ontarian time trend were to be applied, the estimates for the linear time trend model would be precisely the same, with the exception of the time trend (which would impact the results of the BRFM). However, in the case of the GTC models, the resulting estimates differ slightly depending on whether the models include a New Zealand or an Ontarian time trend. [66]

As we explain below, we recommend that the AER discard the LSE models altogether since they fit the data more poorly than the SFA models. However, if the AER decides to use the LSE time-varying models with a GTC time trend, then the AER would need to consider carefully whether the models should incorporate a New Zealand or Ontarian time trend, since this choice (arbitrary as it is) does make a difference to the estimates obtained from the models. The AER should also explain why it thinks its chosen specification is appropriate.

We sought to relax Quantonomics' LSE models via more flexible, nonlinear efficiency trends. However, it is very difficult for LSE models to accommodate efficiency trends as nonlinear as those in SFA models. This is because LSE, as fixed-effect models, are subject to an incidental parameters problem, which broadly means that, as the model becomes more flexible, the number of parameters increases faster than what the sample size can accommodate. Specifically, the fixed effects put a lot of strain onto data requirements for consistent estimation:

- DNSP-specific fixed effects ($\hat{\alpha}_i$) are only consistently estimated when the sample period T is large (i.e., sufficient variation after cross-sectional variation is removed via the fixed effects).
- DNSP-specific coefficients on the linear efficiency trend ($\hat{\delta}_i$) are a particular type of fixed effects. They are only consistently estimated when the sample period T is large.
- Relaxing linearity typically requires an extra set of DNSP-specific coefficients on the efficiency trend term, for instance, each DNSP has three fixed effects coefficients, $(\hat{\alpha}_i, \hat{\delta}_i, \hat{\gamma}_i)$, where $\hat{\delta}_i, \hat{\gamma}_i$ are the coefficients on the efficiency trend terms. The more fixed-effects coefficients involved, the longer the sample period is required for consistent estimation.

65 See Quantonomics, Phase 2 report, pp. 71-72.

66 For brevity we omit the Ontario time trend AJTT models from the reported evaluation. The non-GTC model was equivalent to the NZ time trend specification in terms of model evaluation; the GTC version performed almost identically across the New Zealand and Ontario time trend specifications.

Nevertheless, we relaxed Quantonomics' LSE models by considering a spline specification in the DNSP-specific efficiency trend: a continuous piecewise-linear trend with a kink in 2015. The kink in 2015 is motivated both institutionally (when benchmarking started) and empirically – Quantonomics' LSE residual plots showed a clear kink in residual patterns around 2015.

Specifically, the estimated model is formulated as:

$$lvc_{it} = \beta_0 + \beta_1 ly1_{it} + \beta_2 ly2_{it} + \beta_3 ly3_{it} + \beta_4 lz1_{it} + \beta_5 NZ_i + \beta_5 Ont_i + GTC'\beta_6 + \sum_{j=2}^{13} \alpha_i dj_i + \sum_{j=1}^{13} \delta_i djtm_{it} + \sum_{j=1}^{13} \gamma_i djtp_{it} + \beta_7 nz\_yr_{it} + \varepsilon_{it}$$

Where:

- $dj_i = 1$ if $i = j$ and 0 otherwise.
- $djtm_{it} = \min(t - 2015{,}0)$ if $i = j$ and 0 otherwise.
- $djtp_{it} = \max(t - 2015{,}0)$ if $i = j$ and 0 otherwise.

This results in a continuous time trend in the latent efficiency variable, with a kink in 2015. In 2015 the efficiency variable takes the value of α_i for Australian DNSP i. Prior to 2015 the efficiency variable grows at the rate of δ_i each year, post-2015 the efficiency variable grows at the rate of γ_i each year. We only estimate the GTC time trend variant of this model, which we refer to as LSE-AJTTBREAK-NZ-GTC.[67]

4.5 Summary of models we investigated

Table 3 summarises the models we investigated. The models in grey text are those considered by Quantonomics, which we also considered. The models in the final row of the table are the new/ extended models that we investigated.

[67] We evaluated the specification with the New Zealand time trend added, rather than the Ontario time trend.

Table 3: Models we investigated

BC95		Kumb90		LSE	
Efficiency trend	Cost trend	Efficiency trend	Cost trend	Efficiency trend	Cost trend
JTT (linear)	Linear	JTT (linear)	Linear	ADTT (linear)	Linear
	GTC		GTC		GTC
AJTT (linear)	Linear	AJTT (linear)	Linear	AJTT_NZ (linear)	Linear
	GTC		GTC		GTC
AGEC+JTT: GTC trend for Aus DNSPs	GTC	**AGEC+JTT:** GTC trend for Aus DNSPs	Linear	**AJTT+Ont**	Linear
			GTC		GTC
		AGEC+JTT+ NZ/ONT dummies: - GTC trend for Aus DNSPs - More flexible NZ/Ont trend	Linear	**AJTT (break in 2015)+NZ**	GTC
			GTC		

- JTT: jurisdictional time trend
- AJTT: Aus DNSP specific & jurisdictional time trend
- ADTT: Aus DNSP specific time trend
- ADTT+NZ trend: Aus DNSP specific and NZ time trend
- **AGEC**: step change in efficiency time trend for Aus DNSPs
- **AJTT (break in 2015)**: allows for different continuous time trend before and after 2015
- **GTC:** step change in cost function every 3 years.

Source: Frontier Economics.

5 Evaluation of estimated models

5.1 Selection criteria

We have reviewed the selection criteria applied in section 2.3 of Quantonomics Phase 2 report. We broadly agree with Quantonomics' criteria and commend Quantonomics on developing criteria that may be used to select models in a methodical and consistent manner. It is vital that models are selected for their properties rather than the relative performance of DNSPs under current data. The models that are adopted will likely continue to be applied in the AER's Annual Benchmarking Report, and be used to set DNSP revenue allowances, for some time, and so it is critical that the models can generate reliable estimates as new data are added.

We do however have minor differences of view in the appropriateness of some selection criteria, or how certain criteria should be applied to select models:

- **Reasonableness of output weights:** We agree with Quantonomics that it is useful to consider the sum of the estimated output elasticities, to check if the total output elasticity is close to 1 (i.e., near constant returns to scale). We do not place much reliance on checking the reasonableness of the individual output elasticities on the basis that it is difficult to form accurate priors about individual elasticities. We disagree with Quantonomics that those priors could be informed by output weights "used in the index analysis (which are estimated for Australian DNSPs only), and those derived in opex cost function analysis previously using different data samples or methods."[68] We have explained recently our concerns that the output weights used to construct the MPFP and MTFP indices are unreliable because they have been estimated using a flawed approach. Furthermore, the output weights derived from estimated cost functions in previous benchmarking analyses are also unreliable because those cost functions were mis-specified. That is, after all, the subject of this consultation by the AER. Finally, the high degree of collinearity between some output variables may mean that individual output elasticities are estimated imprecisely.

- **Reasonableness of efficiency and cost time trends:** Quantonomics considers this implicitly by considering how well efficiency scores from each model are correlated to the Opex MPFP index values. We agree that the overall trend in Opex MPFP is reasonable, and estimates should be consistent with that trend. However, we consider this more explicitly by testing whether the pattern of efficiency scores over time appear plausible. By way of example, as explained in section 3.2, we find that the Kumb90 models with linear efficiency time trends to produce implausible patterns of efficiency over time, where efficiency scores tend to converge eventually to 100%. We also consider the estimated time trend of the cost function, and whether the specification is able to reflect non-linear patterns in the cost function.

- **Specification:** Quantonomics considers the *linktest* to evaluate the specification of the SFA models. Unfortunately, this test was implemented incorrectly. When applying the *linktest*, the user should specify the form of the model being tested, so the structure of that model can be adopted in the *linktest* regression. When no such specification is provided by the user after running an *sfpanel* estimation, the *linktest* command by default uses the BC92 time-varying

[68] Quantonomics Phase 2 report, p. 22.

decay model specification to run the test.[69] Since Quantonomics failed to specify the model being tested,[70] each time the *linktest* command was run, the test defaulted to the BC92 specification. This meant, for example, that when Quantonomics tested the specification of a BC95 model, the test regressed observed historical real opex against:

- The fitted values from that BC95 model;
- The square of the fitted values from the BC95 model; but
- *Assuming that the structure of the model being tested was the BC92 model*.

We apply the test correctly by specifying each time the *linktest* command is run the form of the model being tested.

- **Stability to sample changes:** Quantonomics implements this criterion in two ways. Firstly, Quantonomics considers whether the model in question produces efficiency scores differ markedly from previous benchmarking studies, historical Opex MPFP results or established DNSP rankings. We think there is little benefit in doing this, because the econometric models that generated past efficiency scores and DNSPs rankings are clearly mis-specified and are therefore unreliable. If the model in question produces outcomes that differ significantly from past results, there is no way of discerning whether that is because the new model is unreliable, or because the old model was less reliable than the new one. Secondly, Quantonomics tests the stability of model results over different sub-samples. We suggest that model stability should be assessed by adding new data to the sample, and testing how well the model performs with those new data. Specifically, we make an assumption about the true level of efficiency beyond the end of the current benchmarking period (i.e., beyond 2023) and generate opex data for future years consistent with that that level of efficiency. We then fit the model using the expanded data set (i.e., the historical data, plus the 'simulated' future data) to see how well the model estimates the true efficiency level beyond 2023. Such 'out-of-sample' testing is a standard technique for assessing the robustness of models.
- **Parsimony.** We placed only minor emphasis on the criterion of parsimony. All the models we considered were models tested by Quantonomics or involved only minimal extensions to the models evaluated by Quantonomics. We considered that a model satisfied the criterion of parsimony if it converged successfully (in the case of SFA models that needed to be solved numerically), and all relevant parameter estimates could be obtained, using a standard statistical software package, and the models achieve a reasonable goodness-of-fit.

69 Battese, G. E., and T. J. Coelli. 1992. Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. Journal of Productivity Analysis 3: 153–169.

70 This is clear from examination of Quantonomics' Stata output files.

Table 4 below compares the model selection criteria that we have applied against the criteria used by Quantonomics.

Table 4: Comparison of Quantonomics and Frontier Economics model selection criteria

Criterion	Quantonomics	Frontier Economics
Sign/significance of output coefficients	Considered	Considered
Frequency of monotonicity violations	Considered	Considered
Reasonableness of output weights	Considered	Considered (focus on economies of scale)
Coefficient of underground share	Considered	Considered
Reasonableness of efficiency time trend	Considered	Considered
Convergence	Considered	Considered
Goodness-of-fit	Considered	Considered
Performance on specification tests	Considered	Considered
Residual plots	Considered	Considered
Stability to sample changes	Considered (subsets of current data)	Considered (add forecast data)
Consistent efficiency scores	Considered	Considered - reasonableness
Parsimony	Considered (via convergence and goodness-of-fit)	Considered (via convergence and goodness-of-fit)

Source: Frontier Economics.

5.2 Evaluation of models against selection criteria

We now apply the selection criteria to the eighteen time-varying efficiency models considered – five BC95 models, eight Kumb90 models and five LSE models. We also apply the same criteria to the two standard models (the time-invariant inefficiency translog models we denote as LSE-TL and SFA-TL).

Table 5 below summarises the assessment of the models we tested against the selection criteria.

- Those cells shaded green indicate that the model satisfied the criterion fully;
- Those cells shaded amber indicate that the model satisfied the criterion only partially; and
- Those cells shaded red indicate that the model failed to satisfy the criterion.

Details of our evaluation of the models against each criterion are presented in the remainder of this section.

Table 5: Summary of evaluation of models against model criteria

Model	Sign/significance of output coefs	Monotonicity	Reasonableness of weights	Share UGC	Reasonableness of efficiency/cost trend	Convergence
SFA_BC95_JTT_HN						
SFA_BC95_JTT_HN_GTC						
SFA_BC95_AJTT_HN						
SFA_BC95_AJTT_HN_GTC						
SFA_BC95_AGEC_HN_GTC						
SFA_Kumb_JTT_HN						
SFA_Kumb_JTT_HN_GTC						
SFA_Kumb_AJTT_HN						
SFA_Kumb_AJTT_HN_GTC						
SFA_Kumb_AGEC_HN						
SFA_Kumb_AGEC_HN_GTC						
SFA_Kumb_AGECJUR_HN						
SFA_Kumb_AGECJUR_HN_GTC						
LSE_ADTT						
LSE_ADTT_GTC						
LSE_AJTT_NZ						
LSE_AJTT_NZ_GTC						
LSE_AJTTBREAK_NZ_GTC						
LSE_TL (existing model)						
SFA_TL (existing model)						

Model	Goodness of fit	Specification	Residuals	Stability	Consistent efficiency	Parsimony
SFA-BC95-JTT-HN						
SFA-BC95-JTT-HN-GTC						
SFA-BC95-AJTT-HN						
SFA-BC95-AJTT-HN-GTC						
SFA-BC95-AGEC-HN-GTC						
SFA-Kumb-JTT-HN						
SFA-Kumb-JTT-HN-GTC						
SFA-Kumb-AJTT-HN						
SFA-Kumb-AJTT-HN-GTC						
SFA-Kumb-AGEC-HN						
SFA-Kumb-AGEC-HN-GTC						
SFA-Kumb-AGECJUR-HN						
SFA-Kumb-AGECJUR-HN-GTC						
LSE-ADTT						
LSE-ADTT-GTC						
LSE-AJTT-NZ						
LSE-AJTT-NZ-GTC						
LSE-AJTTBREAK-NZ-GTC						
LSE-TL (existing model)						
SFA-TL (existing model)						

Source: Frontier Economics. Note: Some cells in the 'specification' column of the table are unshaded because the linktest command in Stata cannot be applied to the LSE models investigated by Quantonomics.

Sign/significance of output coefficients

A crucial step in the benchmarking process is the estimation of a cost function. The estimated cost function serves two purposes:

- It provides coefficients on output variables to be used in both the roll-forward process and the opex base-step-trend model; and
- It provides estimates of opex consistent with the efficient frontier, allowing comparisons between actual opex and the frontier and thereby allowing efficiency to be estimated.

The coefficients on the key output variables (circuit length, customer numbers and ratcheted maximum demand) can be examined, as well as the significance of the coefficients. It is critical that the three output variables to have a positive impact on opex. The AER regards customer numbers, circuit length and ratcheted maximum demand as well-established drivers of opex. A DNSP's opex is expected to increase as each of these outputs increase. Coefficients for these outputs with a negative sign would not be economically meaningful, suggesting that the estimated cost function is unreliable.

Further, insignificant coefficients might indicate a poor model fit. If these output variables are indeed drivers of opex, then the estimated coefficients should be statistically significant.[71]

Table 6 presents the t-statistics for the evaluated models, providing the sign and significance (a t-statistic greater than 2 is typically considered significant). In all the models all three output variables had a positive coefficient estimate and were statistically significant at the 5% level. With the exception of ratcheted maximum demand in the two AJTT Kumb90 models, the t-statistics were greater than 3, indicating statistical significance at the 1% level.[72,73]

On this basis, we consider that all 18 of the models we tested satisfy this criterion.

Table 6: Output coefficients – t-statistics

Model	t-stat CustNum	t-stat CircLen	t-stat RMDemand
SFA-BC95-JTT-HN	9.76	8.63	11.27
SFA-BC95-JTT-HN-GTC	10.56	8.25	11.29
SFA-BC95-AJTT-HN	11.04	8.22	11.75
SFA-BC95-AJTT-HN-GTC	11.50	7.72	9.80
SFA-BC95-AGEC-HN-GTC	10.43	8.60	11.13
SFA-Kumb-JTT-HN	7.38	6.84	4.60
SFA-Kumb-JTT-HN-GTC	6.91	6.52	4.07

71 Further, insignificant coefficients on output variables would typically indicate that the coefficients have not been precisely estimated.

72 A t-statistic of 3.0 corresponds to a p-value of 0.3%.

73 The RM Demand t-statistics were 2.39 and 2.06 for the SFA-Kumb-AJTT-HN and SFA-Kumb-AJTT-HN-GTC models respectively.

Model	t-stat CustNum	t-stat CircLen	t-stat RMDemand
SFA-Kumb-AJTT-HN	8.00	6.76	2.39
SFA-Kumb-AJTT-HN-GTC	7.94	6.56	2.06
SFA-Kumb-AGEC-HN	6.51	5.82	4.21
SFA-Kumb-AGEC-HN-GTC	6.15	5.32	3.07
SFA-Kumb-AGECJUR-HN	7.64	5.88	3.54
SFA-Kumb-AGECJUR-HN-GTC	7.09	5.85	3.11
LSE-ADTT	4.85	6.84	5.20
LSE-ADTT-GTC	6.05	7.44	4.97
LSE-AJTT-NZ	6.23	6.86	4.23
LSE-AJTT-NZ-GTC	7.45	7.01	3.76
LSE-AJTTBREAK-NZ-GTC	8.17	7.27	3.87
LSE-TL	4.26	6.63	5.14
SFA-TL	3.73	3.04	5.45

Source: Frontier Economics analysis

Frequency of monotonicity violations

For the models considered, the translog versions, we examine the monotonicity of the estimated cost functions for Australian DNSPs. A monotonicity violation indicates that, at the specified level of outputs, an increase in one of the outputs would be associated with a decrease in opex. The estimated cost function would not be considered reliable at that point.

We do however note that the translog cost function, with its flexible specification, can give rise to monotonicity violations, even if specified correctly. Hence, it would be unreasonable to require no monotonicity violations at all. Instead, we consider that a well-specified model should have a low level of monotonicity violations. Specifically, we consider that violations occurring for fewer than 10% of Australian observations are considered acceptable. Violations occurring between 10% and 25% indicates a moderate issue. Violations occurring for greater than 25% of observations indicates a serious issue: 25% corresponds to half of DNSPs having violations for half of observations. Such a model would not be viable for benchmarking purposes.

The frequencies of monotonicity violations for the models tested are provided in Table 7. The AJTT Kumb90 models suffer from excessive monotonicity violations, along with the standard SFA model. These models suffer from serious monotonicity concerns.

Several other models currently exhibit moderate monotonicity concerns: SFA-BC95-JTT-HN, SFA-BC95-AJTT-HN, LSE-ADTT and LSE-TL.

The remaining models either exhibit no monotonicity violations for Australian DNSPs or exhibit a minor level that does not interfere with the application of the estimated models.

Table 7: Frequency of monotonicity violations (Australian DNSPs)

Model	Violations
SFA-BC95-JTT-HN	10.3%
SFA-BC95-JTT-HN-GTC	3.4%
SFA-BC95-AJTT-HN	14.5%
SFA-BC95-AJTT-HN-GTC	0.0%
SFA-BC95-AGEC-HN-GTC	2.6%
SFA-Kumb-JTT-HN	0.0%
SFA-Kumb-JTT-HN-GTC	0.0%
SFA-Kumb-AJTT-HN	55.1%
SFA-Kumb-AJTT-HN-GTC	58.5%
SFA-Kumb-AGEC-HN	0.0%
SFA-Kumb-AGEC-HN-GTC	9.8%
SFA-Kumb-AGECJUR-HN	0.0%
SFA-Kumb-AGECJUR-HN-GTC	0.0%
LSE-ADTT	11.1%
LSE-ADTT-GTC	0.0%
LSE-AJTT-NZ	8.1%
LSE-AJTT-NZ-GTC	8.1%
LSE-AJTTBREAK-NZ-GTC	8.5%
LSE-TL	22.2%
SFA-TL	79.5%

Source: Frontier Economics analysis

Reasonableness of output elasticities

We agree with Quantonomics that output weights should be reasonable. However, while Quantonomics references comparisons with index analysis weights and previous econometric results, we do not consider that consistency with those weights is required:

- We do not view the approach to deriving weights used in index analysis as reliable.[74] Further, the index output weights relate to opex and capital cost rather than solely opex.
- We consider the existing econometric models to be mis-specified. Therefore, previous estimates of output weights derived using those models are unreliable.

We do however agree with Quantonomics that it is reasonable to consider the sum of the estimated output coefficients. We would expect DNSPs to experience moderate returns to scale such that the sum of the estimated coefficients is slightly lower than 1.

We also consider consistency with estimates of well-performing time-varying models. Models that feature relatively extreme coefficient estimates may not be reasonable.

Examining Table 8, estimates appear broadly reasonable, with some exceptions:

- The Kumb90 AJTT models have slight diseconomies of scale. Further, they have relatively low weight on ratcheted maximum demand; and
- The standard SFA model suggests considerable economies of scale: if all outputs increase by 1% opex increases by 0.93%. Further, it has relatively high weight on ratcheted maximum demand.

Table 8: Output coefficients and returns to scale

Model	CustNum	CircLen	RMDemand	RTS
SFA-BC95-JTT-HN	0.386	0.162	0.404	0.953
SFA-BC95-JTT-HN-GTC	0.407	0.154	0.393	0.954
SFA-BC95-AJTT-HN	0.409	0.142	0.406	0.957
SFA-BC95-AJTT-HN-GTC	0.457	0.127	0.376	0.961
SFA-BC95-AGEC-HN-GTC	0.400	0.163	0.390	0.953
SFA-Kumb-JTT-HN	0.487	0.223	0.276	0.987
SFA-Kumb-JTT-HN-GTC	0.503	0.222	0.262	0.987
SFA-Kumb-AJTT-HN	0.616	0.221	0.172	1.009
SFA-Kumb-AJTT-HN-GTC	0.635	0.221	0.153	1.009
SFA-Kumb-AGEC-HN	0.503	0.213	0.282	0.998
SFA-Kumb-AGEC-HN-GTC	0.561	0.205	0.229	0.995
SFA-Kumb-AGECJUR-HN	0.558	0.203	0.233	0.994
SFA-Kumb-AGECJUR-HN-GTC	0.564	0.212	0.217	0.994
LSE-ADTT	0.380	0.230	0.345	0.954

74 See Quantonomics, Nonreliability Output Index Weights ABR25 – Supplementary Analysis, 2025, p.2.

Model	CustNum	CircLen	RMDemand	RTS
LSE-ADTT-GTC	0.432	0.225	0.300	0.957
LSE-AJTT-NZ	0.470	0.216	0.269	0.955
LSE-AJTT-NZ-GTC	0.527	0.208	0.222	0.957
LSE-AJTTBREAK-NZ-GTC	0.544	0.200	0.216	0.959
LSE-TL	0.357	0.238	0.361	0.956
SFA-TL	0.318	0.166	0.443	0.927

Source: Frontier Economics analysis

Coefficient of underground share

As noted by Quantonomics, we expect a higher degree of undergrounding to be associated with lower opex (though higher capital costs).[75] However, unlike the output coefficients, we do not require statistical significance as the relationship need not be precisely estimated since the expected relationship is relatively minor. Further, it is not crucial that the estimated coefficient be precisely estimated for the roll-forward model and base-step-trend calculation given the moderate change in the proportion of undergrounding over time.

A coefficient that is not negative, but statistically insignificant, is acceptable as it can result from imprecise estimates rather than a more severe issue (i.e., mis-specification). A positive and statistically significant coefficient would however raise concerns.

As shown in Table 9 below, the estimated coefficients are negative for the models considered, with the exception of three Kumb90 models:

- SFA-Kumb-JTT-HN-GTC;
- SFA-Kumb-AJTT-HN-GTC; and
- SFA-Kumb-AGEC-HN-GTC.

For these three models the estimated coefficient was 0.01 but statistically insignificant (the largest t-statistic was only 0.42, which implies the estimate was highly insignificant).

Table 9: Coefficient of underground share

Model	Coefficient	t-stat
SFA-BC95-JTT-HN	-0.11	-5.44
SFA-BC95-JTT-HN-GTC	-0.11	-5.59
SFA-BC95-AJTT-HN	-0.15	-9.28
SFA-BC95-AJTT-HN-GTC	-0.15	-8.71
SFA-BC95-AGEC-HN-GTC	-0.10	-5.17

[75] Quantonomics, Phase 2 report, p. 22.

Model	Coefficient	t-stat
SFA-Kumb-JTT-HN	-0.02	-0.54
SFA-Kumb-JTT-HN-GTC	0.01	0.27
SFA-Kumb-AJTT-HN	-0.01	-0.38
SFA-Kumb-AJTT-HN-GTC	0.01	0.42
SFA-Kumb-AGEC-HN	-0.01	-0.36
SFA-Kumb-AGEC-HN-GTC	0.01	0.21
SFA-Kumb-AGECJUR-HN	-0.04	-1.12
SFA-Kumb-AGECJUR-HN-GTC	-0.01	-0.24
LSE-ADTT	-0.10	-3.62
LSE-ADTT-GTC	-0.09	-3.70
LSE-AJTT-NZ	-0.10	-4.10
LSE-AJTT-NZ-GTC	-0.10	-4.33
LSE-AJTTBREAK-NZ-GTC	-0.11	-5.12
LSE-TL	-0.10	-3.40
SFA-TL	-0.12	-2.79

Source: Frontier Economics analysis

Reasonableness of efficiency/cost time trend

We consider that the estimated efficiency time trend from the model should be reasonable and consistent with observed improvements in opex efficiency since approximately 2015, which is when the current opex benchmarking regime was introduced by the AER.

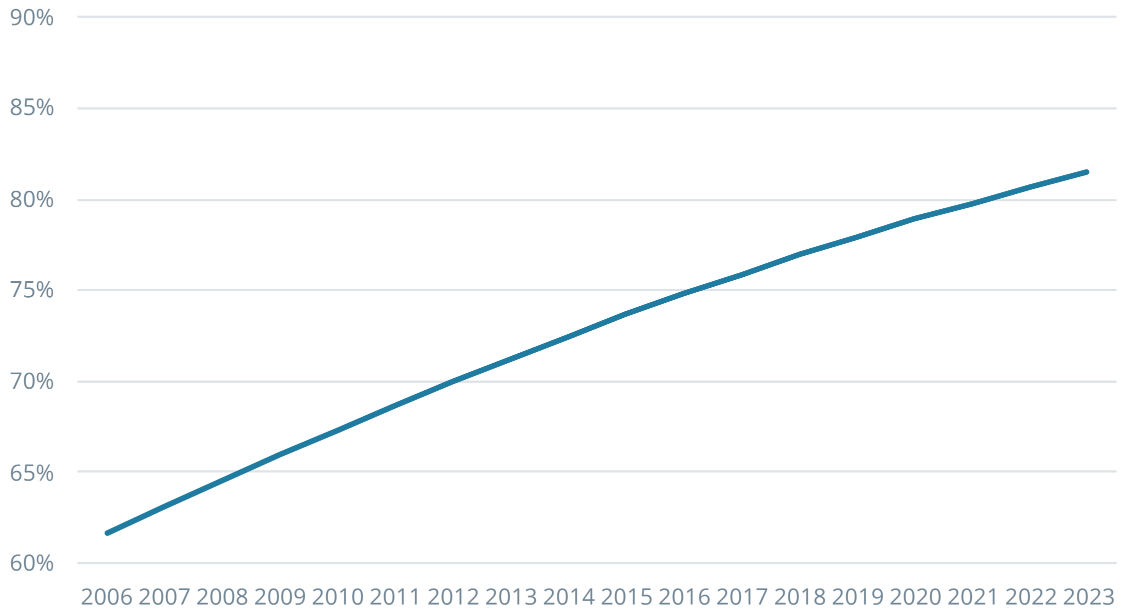
We find that the JTT and AJTT Kumb90 models do not result in reasonable efficiency time trends. In particular, these models exhibit strong convergence for all DNSPs towards 100% efficiency (with the exception of SA Power Networks which converges to a lower efficiency level in AJTT models). While this impacts the individual time-varying estimates of efficiency, the impact on average time-varying efficiency is not so severe. The models do allow an increase in efficiency over time (albeit in a highly restrictive manner) so that the estimates of the cost function and therefore average efficiency over the benchmarking period may be reasonable.

While the LSE time-varying models are restrictive in how they allow efficiency to change over time, they do not suffer the strong efficiency convergence issue exhibited by the JTT and AJTT Kumb90 models.

The standard models fail this criterion due to their assumption of time-invariant efficiency.

While the individual efficiency estimates of the BC95 models exhibit considerable flexibility, and can vary considerably from year to year, we note that the unconditional expected efficiency follows a relatively smooth trend (unlike the highly stylised pattern evident in the JTT and AJTT Kumb90 models). Taking the BC95-JTT-HN estimates, we derive the expected unconditional mean of efficiency for each year from 2006 to 2023. Shown below in Figure 10, the trend is slightly concave, increasing over time but with the increase each year diminishing.

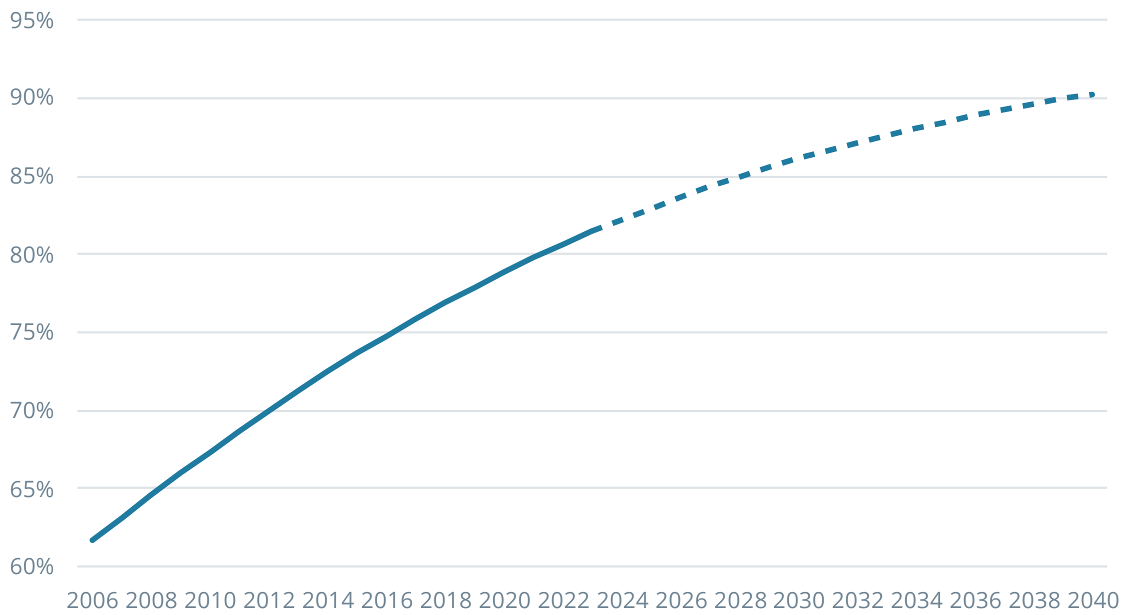
Figure 10: Expected unconditional efficiency of Australian DNSPs - BC95-JTT-HN-GTC-TL



Source: Frontier Economics

We note that the expected efficiency if the BC95-JTT-HN-GTC would converge to 100%. However, this would not occur in the relevant time frame; if we extrapolate the efficiency trend the expected (unconditional) efficiency is only 90% in 2040, illustrated in Figure 11.

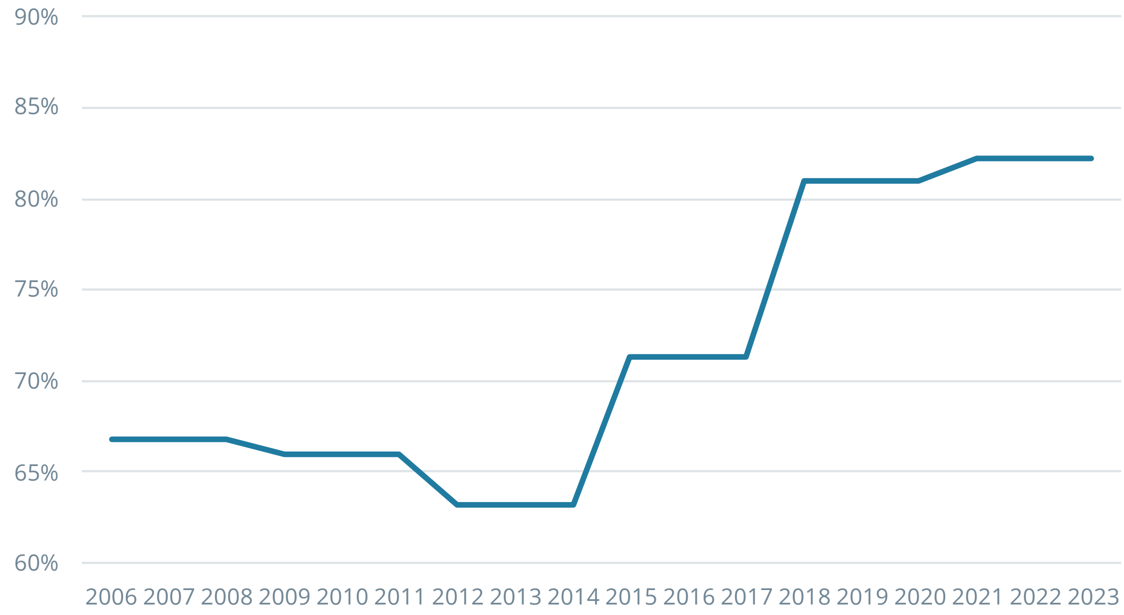
Figure 11: Expected unconditional efficiency of Australian DNSPs - BC95-JTT-HN-GTC-TL (extrapolated)



Source: Frontier Economics

Finally, we consider the trend in the unconditional expected efficiency of the estimated BC95-AGEC-HN-GTC model. The resulting estimates in Figure 12 follow a pattern consistent with the Opex MPFP results, with minor changes in efficiency initially, followed by substantial improvements over the period between 2015 and 2020.

Figure 12: Expected unconditional efficiency of Australian DNSPs – BC95-AGEC-HN-GTC-TL



Source: Frontier Economics

From the perspective of the cost trend, we find that the linear cost time trend models are unsatisfactory, given the observed pattern in the GTC models, with substantial cost rises followed by little change, as demonstrated in **Figure 8**. Using the linear cost time trend could lead to sever mis-specification issues were costs to decline (as per the AER's assumption in the base step trend model of 0.5% productivity improvement per annum.

Convergence

Applying the approach to starting values as set out in section 3.4, all models were able to be estimated having achieved convergence.

We do however note that the standard (time-invariant) SFA model has failed to converge in the 2025 Annual Benchmarking Report. On that basis we consider that the standard SFA model suffers from a convergence issue, insofar that there is a demonstrated issue relevant to estimation moving forward.

Goodness-of-fit

When considering goodness of fit, we first look to the pseudo-adjusted R-squared statistic for each model. In every case, we observe that the SFA models outperform the LSE models. We find that overall, the BC95 models receive slightly higher pseudo-adjusted R-squared statistics when compared to the Kumb90 models.

The worst performing models for each estimation method are the standard models. The Standard SFA model was the lowest scoring of all SFA models, while the Standard LSE model is outperformed by all other LSE models.

Within the BC95 models, the BC95-JTT-HN-GTC and SFA-BC95-AGEC-HN-GTC models perform the best in terms of the pseudo-adjusted R-squared, yet the AJTT models perform better in terms of the BIC.

It should be noted that the BIC of the BC95 models should not be compared to the BIC of the Kumb90 models or the standard SFA model as the log likelihood values are calculated in a different manner.

Table 10: Goodness-of-fit measures

Model	Pseudo-adjusted R2	BIC
SFA-BC95-JTT-HN	99.86%	-416.8
SFA-BC95-JTT-HN-GTC	99.89%	-422.0
SFA-BC95-AJTT-HN	99.68%	-531.1
SFA-BC95-AJTT-HN-GTC	99.79%	-568.3
SFA-BC95-AGEC-HN-GTC	99.89%	-407.1
SFA-Kumb-JTT-HN	99.42%	-1476.9
SFA-Kumb-JTT-HN-GTC	99.40%	-1418.0
SFA-Kumb-AJTT-HN	99.45%	-1456.3
SFA-Kumb-AJTT-HN-GTC	99.43%	-1398.1
SFA-Kumb-AGEC-HN	99.39%	-1402.8
SFA-Kumb-AGEC-HN-GTC	99.40%	-1392.1
SFA-Kumb-AGECJUR-HN	99.42%	-1443.7
SFA-Kumb-AGECJUR-HN-GTC	99.41%	-1401.9
LSE-ADTT	98.25%	
LSE-ADTT-GTC	98.26%	
LSE-AJTT-NZ	98.36%	
LSE-AJTT-NZ-GTC	98.39%	
LSE-AJTTBREAK-NZ-GTC	98.43%	
LSE-TL	98.13%	
SFA-TL	99.10%	-1107.7

Source: Frontier Economics analysis

Performance on specification tests

Quantonomics sought to apply the *linktest* to all the SFA models it investigated. The *linktest* command is not available in Stata for the LSE models used by Quantonomics.[76]

As noted in section 2.3.7, the *linktest* command allows a test of model mis-specification by re-estimating the model being tested, but replacing the cost function explanatory variables with the fitted values and the square of fitted values, keeping other aspects of the model the same. If the square of the fitted values is significant, that is an indication that the model is mis-specified such that the square of the fitted values explains some of the structure of the data.

However, the *linktest* command was not implemented correctly by Quantonomics because Quantonomics did not specify the options used in the original models, for example the specific time-varying specification (BC95 or Kumb90), or the structure of the time varying efficiency.[77] In the absence of these options being specified, the *linktest* command defaulted to evaluating the time-varying BC92 model[78]—a model that was not investigated at all by Quantonomics. In other words, the *linktest* applied by Quantonomics was invalid for the purposes of testing the specified model. This issue can be resolved by providing the *linktest* function with the same model and specification information that is used when initially estimating the model.

We correct the omission, by specifying the model options for each model tested using the *linktest* command. In all the models tested, the squares of the fitted values were found to be statistically insignificant at the 5% level. That is, the *linktest* results provide no evidence of mis-specification. We do however note that that standard time-invariant SFA model has a p-value less than 10%, meaning that the square of the fitted values from that model were statistically significant at the 10% level.

Table 11: Corrected *linktest* results

SFA Model	Z score	p value
SFA-BC95-JTT-HN	0.18	0.859
SFA-BC95-JTT-HN-GTC	0.10	0.924
SFA-BC95-AJTT-HN	0.36	0.718
SFA-BC95-AJTT-HN-GTC	0.40	0.687
SFA-BC95-AGEC-HN-GTC	0.28	0.782
SFA-Kumb-JTT-HN	-0.00	1.000
SFA-Kumb-JTT-HN-GTC	-0.15	0.881
SFA-Kumb-AJTT-HN	-0.16	0.871
SFA-Kumb-AJTT-HN-GTC	-0.04	0.971
SFA-Kumb-AGEC-HN	-0.66	0.512

76 Specifically, the *xtpcse* models.

77 The Stata manual for *linktest* notes that “You should specify the same options with *linktest* that you do with the estimation command” See https://www.stata.com/manuals/rlinktest.pdf.

78 Battese, G. E., and T. J. Coelli. 1992. Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. Journal of Productivity Analysis 3: 153–169.

SFA Model	Z score	p value
SFA-Kumb-AGEC-HN-GTC	-0.09	0.932
SFA-Kumb-AGECJUR-HN	-0.46	0.644
SFA-Kumb-AGECJUR-HN-GTC	-0.35	0.727
SFA-TL	-1.90	0.057

Residual plots

The residual plot of a well-specified model should ideally demonstrate that the residuals are randomly distributed across both time and magnitude, with a mean of zero. This indicates that the model is not systematically misestimating true values in a predictable manner, indicating that the model is not adequately explaining the full structure of the data. This may be shown through residuals consistently increasing or decreasing over time, or residuals increasing or decreasing as the fitted values increase or decrease.

Of the SFA models, only the Kumb-JTT-HN and Kumb-JTT-HN-GTC produced residuals plots that raised concern. Some structure was visible in the residual plots of other SFA models, but the patterns were not sufficiently strong to conclude that the models were obviously mis-specified. By contrast, for the LSE models only the residual plot of the AJTTBREAK-NZ-GTC did not display strong patterns. For both of the standard models, the residual plots demonstrated significant and apparent structure. As such, they were considered to fail this test.

The residual plots for each model tested are presented in Appendix C.

Stability to sample changes

An important consideration for these models is the capacity to fit and explain new data without producing nonsensical results. The increasing monotonicity violations of the AER's existing models as new data was added was, after all, one of the key motivations for this review. In short, a model needs to be stable to changes in sample and in sample size.

Quantonomics evaluated this by dividing the existing sample into numerous sub-periods and testing the ability of models to incorporate increasingly larger subsets of the original data. In our view this approach does not sufficiently test the stability of model results to sample changes. We already know that the standard models are capable of producing reasonable-looking estimates for subsets of the historical data, even though the underlying model is mis-specified, since the incidence of statistical problems (e.g., monotonicity violations) have become more severe over time. As such, testing the capacity of the models to use subsets of old data (in effect, testing what the model would have outputted had it been used in 2019, for instance) does not address the primary concern.

Instead, we use the baseline BC95-JTT model to generate additional years of data such that the true efficiency scores of each DNSP are held constant beyond 2023. We then apply each model to the new dataset (included the post-2023 simulated data) and examine whether the models produce sensible results when considering the known underlying efficiency scores. If a model continues to predict increasing efficiency scores over the generated period (when in fact the true level of efficiency remained constant), that is an indication that the model is incapable of incorporating additional data well. If the model predicts plateauing efficiency scores, it is an indication that the model has successfully determined that efficiency has ceased to improve and is capable of fitting the new data well.

Of the tested models, we found that two BC95 models with the AJTT specification failed to fit the generated data well, because those models estimated that efficiency continued to improve beyond 2023, even though, by construction, efficiency post-2023 remained constant. Similarly, we found that both Kumb90 models with the JTT specification were unable to fit the generated data well, as those models estimated that efficiency continued to improve beyond 2023.

We found that none of the time-varying LSE models were able to fit the generated data well as all those models estimated that efficiency continued to improve beyond 2023.

We present plots that extend the efficiency trends of each model out to 2030 in Appendix C, with the period of generated data indicated by a dotted line at 2023, denoting the final year of the benchmarking period using the 2006 to 2023 dataset.

Consistent efficiency scores

A key outcome of the benchmarking process is the ability to estimate efficiency of DNSPs via the resulting efficiency scores. Models that generate unreasonable or implausible efficiency scores are unlikely to be reliable.

We place little emphasis on consistency with previous benchmarking studies and Opex MPFP results. Our view is that previous benchmarking studies were based on mis-specified models.[79]

The standard SFA model has performed poorly recently, with very low estimates appearing for some DNSPs. In particular, Ausgrid has an implausibly low efficiency score estimate less than 30% using 2006-23 data. On this basis alone such models should be excluded.

While most models result in average efficiencies for the Australian DNSPs (over the period 2006 to 2023) of between 70% and 74%, several other models have average scores that somewhat higher or lower than this range:

- SFA-BC95-AJTT-HN and SFA-BC95-AJTT-HN-GTC have relatively high average efficiency scores across the Australian DNSPs, averaging over 81% across 2006-2023; and
- SFA-Kumb-AGEC-HN and SFA-Kumb-AGEC-HN-GTC similarly have relatively low average efficiency scores, averaging 60% and 64% respectively.

While it remains plausible that such estimates are valid, further investigation would be required before inclusion of these models since the differences in average efficiency estimates are material.

Parsimony

As noted by Quantonomics, “if two models perform similarly, the more parsimonious model is usually preferred.” Further, the “preferred model should also be straightforward to apply”.[80]

During the analysis we found that all models were parsimonious in the sense that estimation was straightforward: established models estimated using an available Stata software package. Further, results were easily obtained in little time.

None of the models were excessively complex, in the sense that each model involved a relatively small modification or variation to the standard models.

On this basis, we consider that all estimated models satisfy the parsimony criterion.

79 The time-invariant restriction may have been appropriate in the earliest studies, but the efficiency scores from these studies bear little relevance to recent efficiency estimates.

80 Quantonomics Phase 2 report, p. 24.

5.3 Cobb-Douglas vs. translog specifications

The Cobb-Douglas specification of the econometric cost functions are a special case of the translog specifications with a less flexible form. Quantonomics has previously undertaken statistical tests of the Cobb-Douglas versus the translog specifications, where null hypothesis for this test is that the restrictions imposed on the translog model to obtain the Cobb-Douglas are consistent with the data. These tests soundly reject the Cobb-Douglas simplification of the translog model in all cases that Quantonomics investigated.[81]

We repeated those statistical tests to assess whether there is any evidence that Cobb-Douglas versions of the 18 models we investigated are consistent data explain the available data. Specifically, we apply the likelihood ratio test for SFA models, and the Wald for the LSE models, as Quantonomics has previously done. In all cases, we use data over the long sample period. The likelihood ratio considers the improvement in the log-likelihood (the measure of goodness of fit) when we relax the Cobb-Douglas restriction and estimate the translog model, the Chi-squared values of the LSE models have a similar interpretation. The accompanying p-values reflect the level of significance of the resulting test statistic, with values less than 5% taken as statistically significant evidence that the Cobb-Douglas model should be rejected in favour of the translog model. We find that the Cobb-Douglas model is rejected in 17 of the 18 models estimated, as shown in Table 12.

Table 12: Tests of Cobb-Douglas vs translog models

Model	LR	Chi-sq	p-value
SFA-BC95-JTT-HN	117.96		0.0%
SFA-BC95-JTT-HN-GTC	122.25		0.0%
SFA-BC95-AJTT-HN	110.53		0.0%
SFA-BC95-AJTT-HN-GTC	111.06		0.0%
SFA-BC95-AGEC-HN-GTC	124.7		0.0%
SFA-Kumb-JTT-HN	15.83		1.5%
SFA-Kumb-JTT-HN-GTC	18.76		0.5%
SFA-Kumb-AJTT-HN	28.47		0.0%
SFA-Kumb-AJTT-HN-GTC	30.42		0.0%
SFA-Kumb-AGEC-HN	11.78		6.7%
SFA-Kumb-AGEC-HN-GTC	14.07		2.9%
SFA-Kumb-AGECJUR-HN	13.15		4.1%
SFA-Kumb-AGECJUR-HN-GTC	16.03		1.4%

81 Frontier Economics, Benchmarking analysis of Energex's and Ergon Energy's opex, January 2024, Table 2.

Model	LR	Chi-sq	p-value
LSE-ADTT		47.07	0.0%
LSE-ADTT-GTC		52.45	0.0%
LSE-AJTT-NZ		53.07	0.0%
LSE-AJTT-NZ-GTC		57.71	0.0%
LSE-AJTTBREAK-NZ-GTC		68.45	0.0%

Source: Frontier Economics analysis

These test results suggest that the translog models fit the data significantly better than the Cobb-Douglas models. We discuss the implications of this finding for the models that the AER should adopt in its benchmarking analysis in the next section.

5.4 Recommended models

Our first key finding is that the SFA and LSE models that the AER currently uses perform the most poorly of all the models we assessed against the model selection criteria. There is now overwhelming evidence that the time-invariant inefficiency models that the AER has hitherto relied on are badly mis-specified and should be discontinued altogether.

The key reason why these models fail is because of their constant efficiency assumption, which is inconsistent with the available data. The models are incapable of accounting for the fact that at least some Australian DNSPs have become materially more efficient over time, particularly since the AER introduced its current benchmarking regime in 2015.

Of the 18 time-varying inefficiency models we tested, we identified two BC95 models and one Kumb90 model that satisfied all the selection criteria:

- SFA-BC95-JTT-HN-GTC;
- SFA-BC95-AGEC-HN-GTC; and
- SFA-Kumb-AGECJUR-HN-GTC.

Given that these are clearly the best performing of the models tested, we recommend that these models replace the existing benchmarking models the AER currently employs, and that all three models receive equal weight by default.

Quantonomics makes the point that there is considerable uncertainty over the estimates of efficiency produced by the econometric benchmarking models and suggests that there may be some benefit to using differently specified models as this may improve the reliability of the estimated efficiency scores. We agree with the general proposition that one legitimate way to deal with model uncertainty is to use estimates from a range of different models, provided that the none of the models are clearly inferior to the others, and do not produce obviously biased estimates of efficiency of the cost function parameters.

Contrary to Quantonomics' findings, we recommend that the AER abandon the LSE specification altogether. As the analysis above shows, the time-varying inefficiency LSE model performed unambiguously worse than the three models identified above. The LSE models are highly restrictive in their specification and fit the data more poorly than the SFA models. We think that given the statistical evidence, there is no longer a case to continue using the LSE models.

Finally, we recommend that the AER give primacy to the translog specification over the Cobb-Douglas specification. Quantonomics has previously presented compelling evidence that the translog specification fits the data much better than the Cobb-Douglas versions of the same models. However, it was useful to the AER to retain the Cobb-Douglas versions because in some cases the translog models (which are highly flexible) suffered from monotonicity violations so severe that they could not be used reliably. In those instances, the availability of the Cobb-Douglas models meant the AER could still obtain efficiency scores and cost function estimates with which to perform the benchmarking analysis. However, as we have shown above, the two BC95 models and the Kumb90 model, suffer from very minor or no monotonicity violations (when estimated using historical data over the period 2006 to 2023).

Therefore, we propose that, where possible, the AER should rely only on results from translog versions of the models recommended above and use results from Cobb-Douglas versions of those models only if the translog versions exhibit clear statistical problems (e.g., severe monotonicity violations).

6 Estimation of efficient base year opex

6.1 Options considered

As set out in Section 2.5, Quantonomics considered that the current benchmarking roll-forward approach could potentially be replaced with an approach using the final year of the benchmarking period directly to estimate efficient base year opex (in the case where the base year is the final year of the sample), or by reducing the use of the roll-forward approach to bridge between the final year and base year used to forecast efficient opex over the next regulatory period.

6.1.1 Roll forward from sample average

Under the AER's current approach, the sample average estimate of efficient opex is rolled forward, using the estimated cost function, to obtain an estimate of efficient opex in the base year.

The sample average opex actually incurred by the DNPS and sample average efficiency is examined. If the AER finds that the DNSP's average efficiency score over the benchmarking period is lower than the OEF-adjusted efficiency target, then the AER makes an adjustment to the sample average opex equal to difference between the DNSP's estimated efficiency score and the OEF-adjusted efficiency target. This 'adjusted' level of opex is interpreted as an estimate of efficient opex in the middle of the benchmarking period. That figure is then rolled forward to the base year accounting for growth in the DNSP's outputs and the impact of technology (i.e., trends in real opex over time). The estimated cost function coefficients are used to derive the impact of output growth and technology.

The same approach can be readily applied in a time-varying efficiency context. The average efficiency over the sample period can be obtained as the average of the efficiency estimates for the individual years in the sample and then used to obtain the starting point for the roll-forward procedure.

6.1.2 Roll forward from final year

Quantonomics notes that, if time-varying inefficiency models are adopted, then the AER can obtain an efficiency score in individual years of the benchmarking period—including the final year. If the final year happens to coincide with the DNSP's base year, then the estimated efficiency score in that year can be used directly to derive an estimate of efficient base year opex.

If the final year of the benchmarking period is not the base year, then the estimated efficient opex in the final year of the period can be rolled forward (using a similar approach to the existing roll-forward method) to obtain an estimate of efficient base year opex.

This would reduce the reliance on the estimated elasticities in deriving the impact of output growth on opex: output growth would typically be either zero (if the final year is the base year) or a single year's growth (if the base year were one year after the final year of the benchmarking period).[82]

[82] In the case of translog models, the applied elasticities are those at the sample average rather than the Australia average or that of the DNSP.

6.2 Feasibility of models for final year adjustment

Implicit in the final year approach is the requirement that one can adjust the final year opex using the final year estimated efficiency to obtain an estimate of efficient opex in the final year. This requirement is not satisfied for either the Kumb90 or LSE time-varying models.

The base Kumb90 model essentially estimates a single efficiency for the DNSP and then applies a stylised time-varying adjustment (common to all Australian DNSPs) to best match the data. An increase in the final year actual opex will have almost no impact on the estimated final year efficiency, except for the extent to which it reduces the DNSP's sample average efficiency.[83] As it is one year of approximately twenty for the long sample, this impact will generally be very minor.

Consider the base translog Kumb90 model (SFA-Kumb-JTT-HN-TL). In 2023 Evoenergy had an opex of $71,602,000 and a resulting final year efficiency of 88.7%. If we were to substitute opex of $81,602,000 for 2023, the resulting final year efficiency would reduce to 88.5%. This is an immaterial reduction and implies that opex at 75% efficiency is $96,323,00 rather than $84,645,000 using the actual opex.[84]

This would mean that estimates of efficient base year opex increase materially as final year opex increases, almost one-for-one. This results in:

- Estimates of efficient base year opex that unreliable; and
- Potentially strong incentives in the final year to increase the estimate of efficient base year opex thereby avoiding base year reductions.

This consideration applies to other Kumb90 models, such as the AJTT models, as there is similarly a limited impact of the final year opex on the estimated final year efficiency. For similar reasons this also applies to the time-varying LSE models.

Only the BC95 models avoid this inflexibility of estimated final year efficiency. Under these models the estimated final year efficiency evaluates the final year opex against the half-normal inefficiency distribution and estimates of the variance parameters. The estimated efficiency in the final year will fully reflect the comparison between the actual final year opex and the final year frontier opex, providing the expected efficiency given the two opex values.[85]

Consider the base translog BC95 model (SFA-BC95-JTT-HN-TL). In 2023 Evoenergy had an opex of $71,602,000 and a resulting final year efficiency of 76.9%. If we were to substitute opex of $81,602,000 for 2023, the resulting final year efficiency would reduce to 68.5%. This is a material reduction.

The key point is that Quantonomics' suggestion of using efficiency score estimates in the final year of the benchmarking period are feasible only in the case of the BC95 models; this approach cannot be used in the case of the Kumb90 and time-varying LSE models because, under those models, the estimate of efficiency in a particular year is not reflective of, or sensitive to, the level of opex actually incurred by the DNSP in that year.

6.3 Asymmetric final year adjustment & EBSS incentives

Under the AER's current approach, the sample average opex may be adjusted downwards to account for any difference between the target efficiency score—i.e., 75% adjusting for OEFs, and the period average efficiency score:

[83] And potential minor impacts on the Australian time varying adjustment factor.

[84] Implied opex at 75% efficiency calculated as 71,602×(88.7%/75%) and 81,602× (88.5%/75%).

[85] Using the Battese-Coelli (1988) expression for estimated efficiency.

$$C_{PA}^{e} = C_{PA} \times \min\left\{\frac{\theta_{PA}}{75\%/(1+OEF\%)}, 1\right\},$$

where C_{PA}^{e} is the efficient period-average opex, C_{PA} is the real period-average opex, θ_{PA} is the period-average estimate efficiency score and $OEF\%$ is the OEF adjustment to the DNSP's efficiency target. This provides an estimate of the average opex over the period if the DNSP were not materially inefficient.

This approach is asymmetric, in that DNSPs with efficiency below the target have their sample average opex adjusted downwards, where DNSPs with efficiency above the target do not have their sample average opex adjusted upwards. For example, the estimate of efficient opex of a DNSP with an efficiency score of 95% is not increased to reflect the opex of a DNSP with an efficiency score of 75%. Instead, the actual average opex would be used as the efficient period-average opex (with the implicit efficiency of 95%).

This asymmetry has historically been of little concern as the base year opex of DNSPs with an average efficiency score over the benchmarking period above 75% is deemed to be efficient; in such instances, the AER does not perform the roll-forward calculation from the middle of the benchmarking period to the base year, to derive an estimate of efficient base year opex.

We now examine the implications of applying such an asymmetric approach to the final year approach proposed by Quantonomics, without exempting DNSPs that are at least 75% efficient in the final year, in particular the typical circumstance of the base year being subsequent to the final year of the benchmarking period.[86]

Consider a final year opex (e.g., $100) which reflects a status quo, that results in the DNSP matching the target efficiency (i.e., 75% adjusted for OEFs) in the final year. At this level of opex, the estimate of efficient final year opex is equal to actual final year opex ($100). The estimate of efficient base year opex will therefore equal the actual final year opex with some small adjustments for output change etc. between the final year and the base year. Assuming no improvements in efficiency, we would expect that actual base year opex will exceed the estimate of efficient base year opex approximately 50% of the time, due to changes in the idiosyncratic noise term.

If the DNSP were to reduce the final year opex further (e.g., $90), its estimated efficiency score would increase (83.3%) but the estimate of efficient final year opex will be equal to the (reduced) actual final year opex ($90). The estimate of efficient base year opex will accordingly reduce. Assuming no improvements in efficiency, we would expect that actual base year opex will exceed the estimate of efficient base year opex more than 50% of the time, due to changes in the idiosyncratic noise term.

That is, in the absence of efficiency improvements between the final year and the base year, we would expect the DNSP to be found to be materially inefficient more than 50% of the time, regardless of the DNSP's actual efficiency. This is not a reasonable outcome.

A key issue with the asymmetric approach is that any efficiency improvements made in the final year are implicitly assumed to be incorporated into the baseline for efficient base year opex. Any savings made need to be maintained and in practice added to in order not to be found to be materially inefficient. This gives rise to a strong incentive to delay efficiency gains until after the final year of the benchmarking period, to avoid a base year adjustment.

86 Typically, though not always, the penultimate year of a regulatory period is used as the base year for forecasting efficient opex over the next regulatory period. The penultimate year of a regulatory period is not available for inclusion within the benchmarking sample period used by the AER in its benchmarking analysis, when making a revenue determination.

In the standard approach, DNSPs are incentivised to make efficiency gains whenever the opportunity arises, including in the base year. While the opex allowance will be lower as a consequence, the DNSP is compensated via the Efficiency Benefit Sharing Scheme (EBSS).

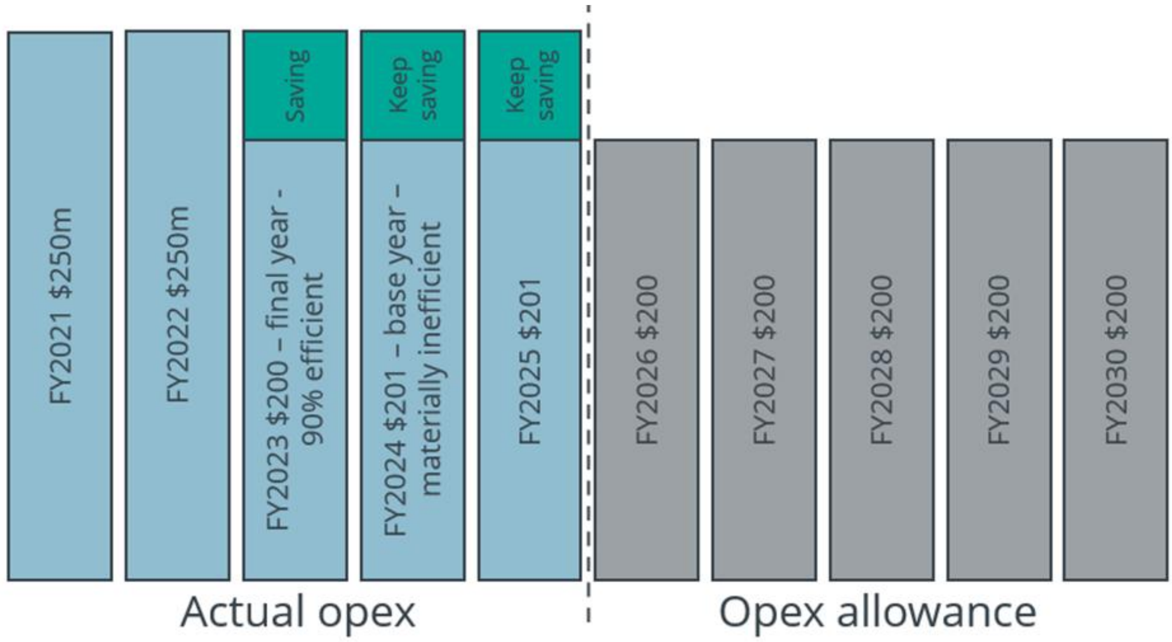
Under the asymmetric final year approach, a DNSP that reduces the final year opex beyond the target efficiency level will lower the estimate of efficient base year opex and therefore the opex allowance (assuming the DNSP is unable to improve efficiency and therefore is subject to a base year reduction). However, in such a circumstance the EBSS would (under the AER's existing framework) no longer be applied (due to the base year reduction) and the DNSP would not receive the EBSS benefit of the efficiency improvements in the final year.

This consideration creates a powerful incentive for DNSPs to defer efficiency improvements until the base year, delaying the benefit customers would receive from efficiency improvements.

By way of illustration, consider the following example, set out in Figure 13. A DNSP has the opportunity to realise considerable permanent savings in the third year of a regulatory period, reducing its opex from $250 (the allowance for each year of the regulatory period) to $200, a saving of $50, resulting in an efficiency score estimate of 90% for that year (which coincides with the final year of the benchmarking period). If the AER were to apply the asymmetric final year approach, then the revealed opex of $200 in year 3 would be rolled forward to estimate the efficient opex in the base year. Suppose, for simplicity, that the rate of output growth and the rate of technical change are both zero, such that the rolled-forward estimate of efficient opex in the base year is $200.

If the DNSP expects that its opex in the base year will be $201, it would anticipate that opex in the base year would be found to be materially inefficient due to the $1 efficiency loss. If that occurs, under the AER's current approach the EBSS would not apply. This would mean that the DNSP would maintain the saving for the remainder of the regulatory period, but the EBSS benefits associated with the original $50 saving would not be received over the first three years of the next regulatory period (as would be the case if the EBSS had continued to apply).

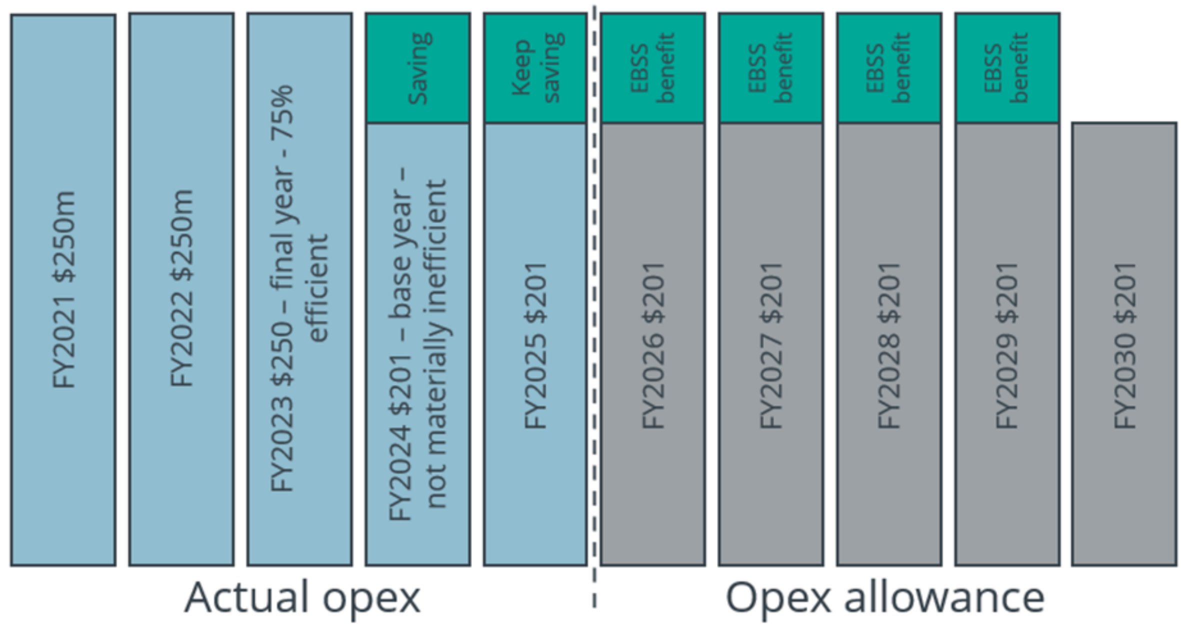
Figure 13: Allowance, opex and EBSS benefits under asymmetric final year approach



Source: Frontier Economics

However, if the DNSP were to defer the saving until year four, it would be found to be not materially inefficient, thereby receiving an EBSS benefit for four years of the next regulatory period, as shown in Figure 14 below. Thus, the DNSP would have a strong incentive to defer efficiency gains so as to avoid a base year opex reduction. By deferring the efficiency gain by a single year, it would be able to keep the benefits of the efficiency gain for an additional three years that would otherwise be forfeit if it made the savings as soon as the opportunity arose in year 3.

Figure 14: Allowance, opex and EBSS benefits – defer gain



Source: Frontier Economics

The key problem we identify above is that under an asymmetric approach to assessing the efficiency of base year opex, using efficiency estimates in the final year of the benchmarking period, DNSPs that can make material efficiency improvements in that final year of the benchmarking period may face strong incentives to defer those savings to maximise their EBSS rewards. The DNSP will always be able to anticipate the final year of the benchmarking period (even if the base year is uncertain), so delaying efficiency gains in this way would be a viable strategy. However, this means that the benefits to consumers of such efficiency gains would also be deferred, instead of being realised when opportunities for efficiency become available to the DNSP—a principle at odds with the intent of the EBSS.

One way to overcome this incentive problem would be to apply a symmetric approach when assessing the efficiency of base year opex. That is:

- If the DNSP's estimated efficiency in the final year is below the efficiency target, then the DNSP's actual opex in that year would be adjusted down to a level consistent with the efficiency target (as the AER currently does, when evaluating the efficiency of a DNSP's average opex over the benchmarking period); and
- If the DNSP's estimated efficiency in the final year is above the efficiency target, then the DNSP's efficient opex in the final year would be determined by adjusting its actual opex in that year up to a level that is consistent with the efficiency target.

The symmetric adjustment would be implemented as follows:

$$C_T^e = C_T \times \frac{\theta_T}{75\%/(1 + OEF\%)},$$

where C_T^e is the DNSP's efficient opex in the final year of the benchmarking period T, C_T is DNSP's actual opex in year T, θ_T is the DNSP's estimated efficiency score in year T and $OEF\%$ is the OEF adjustment to the DNSP's efficiency target.

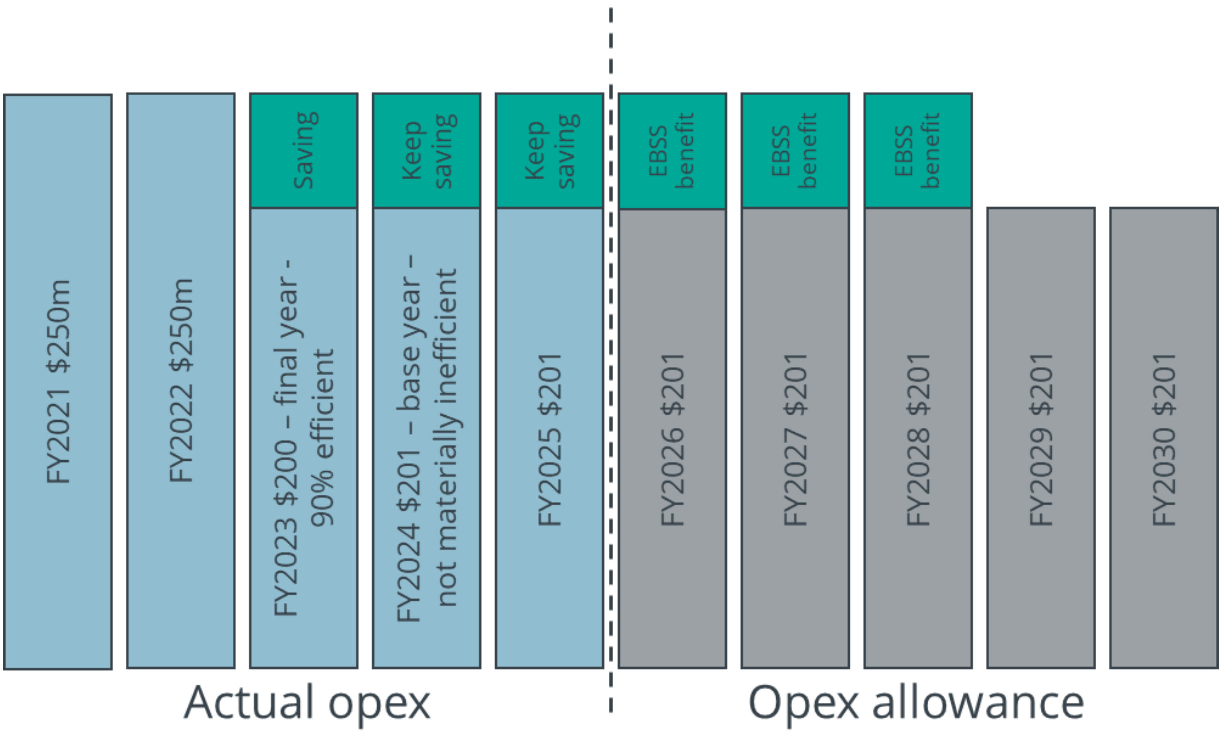
That is, if the DNSP's estimated efficiency score in the final year is higher than its efficiency target, the DNSP's estimate of efficient final year opex (for the purposes of determining the efficient level of base year opex) would greater than its actual opex in the final year of the benchmarking period. This provides an estimate of final year opex that would be considered as efficient. This removes the distortion of incentives caused by the asymmetric approach, which resulted in the incentive to defer efficiency improvements.

To see this, we return to the stylised example provided above. Once again, assume that in the first and second years of the regulatory period the DNSP incurs opex of $250, but has the opportunity to make a permanent saving of $50 in the third year. Realising the saving would result in an efficiency score in the final year of 90% and, as before, its efficiency target is 75% (i.e., the OEF adjustment is zero). Under the symmetric approach, the estimate of efficient opex in the third year would be:

$$\$240 = \$200 \times \frac{90\%}{75\%}.$$

Assume further, as before, that the rate of output growth and the rate of technical change are zero, such that the rolled-forward estimate of efficient opex in the fourth year of the period (i.e., the base year) is $240. In these circumstances, the DNSP would expect that it would not face any base year opex reduction, even if its actual opex in the base year turned out to $201. As a result, the DNSP would be incentivised to make the saving in the third year of the regulatory period, when the opportunity to do so arises, because the DNSP would expect to retain the benefit of those savings for six years (i.e., the three remaining years of the current regulatory period, and during the first three years of the next period via EBSS payments), as shown in Figure 15.

Figure 15: Allowance, opex and EBSS benefits under a symmetric final year approach



Source: Frontier Economics

The key takeaway is that if the AER wishes to assess the efficiency of DNSP's opex using the estimated efficiency score in the final year, we recommend that the AER estimate the efficient level of opex in the final year in a symmetric manner, by calculating the efficient level of opex consistent with the efficiency target for that DNSP.

6.4 Operating environment factors

Under the AER's approach, Operating Environment Factors (OEFs) reflect differences in the environment in which DNSPs operate which have material impact on cost yet are not included in as explanatory variables in the econometric benchmarking models. These important uncontrollable cost differences are accounted for by adjusting the 75% efficiency target. The adoption of time-varying inefficiency models may require some modification to the way OEF adjustments are calculated, depending on whether:

- the AER maintains its current approach of rolling forward estimates of efficient opex from the middle of the benchmarking period; or
- the AER decides to derive estimates of efficient base-year opex using the estimated efficiency score in the final year of the benchmarking period.

6.4.1 OEFs under the sample average approach

To calculate the required OEF adjustment under the current approach, the AER performs the following steps for each individual OEF (for example the OEF adjustment for ownership of sub-transmission assets):

- Estimate the average real cost of the relevant OEF for each DNSP over the sample period;
- Compare the OEF real cost to the 'efficient base opex' – the average real opex over the benchmarking period, applying a reduction to obtain 75% efficiency if the average efficiency across models is below 75% over the sample period. This expresses the OEF as a percentage of 'ideal' opex;
- Calculate the OEF reference point as the average OEF (as a percentage) of the reference DNSPs, weighting by average customer numbers over the sample period; and
- Calculate the OEF adjustment by comparing the DNSP's OEF and the average reference DNSP OEF (both expressed as a percentage).

This approach could be adopted under an approach with time-varying efficiency. The only step that would be impacted is the calculation of efficient base year opex. For each model the average efficiency over across the sample period would be calculated, rather than simply taking the single efficiency estimate for the model.

The other steps of the approach would require no change.

6.4.2 OEFs under the final year approach

In general, the OEF adjustments required under a final year approach would be different to those required under a sample average approach in that the OEF adjustment should specifically reflect the OEF costs in the final year. This gives rise to the following considerations:

- The OEF cost would in principle need to reflect the estimated cost in the final year. We note that some OEFs are time varying, for example Ergon Energy's cyclones OEF. However, in some circumstances a sample average may be taken as an estimate of the OEF in the final year due to challenges in estimating the OEF.

- It is unclear whether reference DNSPs would be those that average at least 75% efficiency over the sample period, or those that average at least 75% in the final year. We note that DNSPs may have efficiency greater than 75% in the final year yet have an average below 75% over a sample period. It is also in principle possible for a DNSP to average above 75% over a sample period but fall below 75% in the final year (e.g., if its opex in the final year is unusually high).

6.5 Should the AER roll forward from the sample average or use estimated efficiency in the final year to estimate efficiency base year opex?

If the AER were to adopt the time-varying inefficiency models, the current approach using the benchmarking roll-forward model from the middle of the benchmarking period (rather than using final year estimates of efficiency to estimate efficient base year opex) may still be applied, but applying the average time-varying efficiency estimates over the benchmarking period to estimate the average efficient opex over the period. The process is familiar to DNSPs and the Quantonomics report does not suggest that a change is needed. Furthermore, no change would be required to the way the AER calculates OEF adjustments. For these reasons, we recommend that the AER retain its current approach of rolling forward estimates of efficient opex from the middle of the benchmarking period.

In our view, the AER should consult with stakeholders comprehensively before changing its approach to estimating efficient base year opex, to ensure there are no unanticipated consequences of such an approach. There are a number of complex issues that the AER would need to take into account, and address, before changing its current approach of rolling forward from the middle of the benchmarking period:

- Not all estimates of efficiency in the final year are valid for use in directly estimating efficient final year opex. As noted above, the Kumb90 models and LSE models generate time-varying efficiency estimates that are substantially anchored to the estimated relationship and so will not adjust to fully reflect actual opex in the final year. Thus, only BC95 models generate estimated efficiencies in the final year that could potentially be used to estimate efficient final year opex.
- If the AER decides to roll forward estimates of efficient opex using estimated efficiency scores in the final year of the benchmarking period, the AER must do so in a way that does not undermine incentives for efficiency by removing EBSS rewards from those DNSPs that could realise large savings in the final year of the benchmarking period.

A Approach to selecting starting values

This Appendix sets out the approach taken by Frontier Economics to setting starting values for time-varying SFA models using the *sfpanel* package in Stata.

Step 1 – estimate the basic Cobb-Douglas cost function using least squares regression

This step serves to generate an approximate cost function to be used in estimating SFA models, thereby providing those models with a reasonable starting point for the cost function. The simple least squares regression is used in Stata, with the following equation estimated for linear cost time trend models:

$$\mathrm{lvc_{it}} = \beta_1 + \beta_2 \log(custnum_{it}) + \beta_3 \log(circlen_{it}) + \beta_4 \log(rmdem_{it}) + \beta_5 \log(shareugc_{it}) + \beta_6 t + \beta_7 NZ_i + \beta_8 Ont_i$$

The following equation is estimated for GTC models:

$$\mathrm{lvc_{it}} = \beta_1 + \beta_2 \log(custnum_{it}) + \beta_3 \log(circlen_{it}) + \beta_4 \log(rmdem_{it}) + \beta_5 \log(shareugc_{it}) + GTC_t' \beta_{GTC} + \beta_7 NZ_i + \beta_8 Ont_i$$

The resulting coefficient estimates are used as starting values for the cost function in the subsequent SFA regression. In addition, the standard error of the regression is used as a starting value for the total inefficiency/noise variance in the subsequent SFA estimation.

Step 2 – estimate the Cobb-Douglas time-invariant efficiency half-normal SFA model

Using the estimated cost function from step 1 as starting values, along with the error variance estimate, with values of 0.1 for the relative size of inefficiency and noise, the *xtfrontier* package is used to estimate the time-invariant SFA model.[87,88]

The SFA model is restricted to have a half-normal distribution of inefficiency, by setting μ to equal zero. The resulting model is relatively straightforward to estimate, as appropriate starting values have been set, the cost function is the relatively simple Cobb-Douglas model, and the inefficiency distribution is set to half-normal.[89]

The resulting SFA model yields three key outputs, to be used as starting values in the subsequent time-varying SFA model:

- Estimates of the Cobb-Douglas cost function
- Estimates of the variance of the inefficiency component
- Estimates of the variance of the idiosyncratic error component

Step 3 – estimate the Cobb-Douglas time-varying efficiency SFA model

The time-varying efficiency SFA model can then be estimated. Using the *sfpanel* package, four starting value vectors are to be provided:

- The frontier cost function – taken as the cost function estimated in Step 2

[87] This value of 0.1 has historically been used by Quantonomics as a starting value for that component.

[88] As the model is a half-normal model, any starting value of μ is ignored as μ is restricted to equal zero.

[89] See Section 2.2.1 for an overview of the half-normal restriction.

- The distribution of inefficiency. This is taken as:
  - The log of estimated σ_u^2 for BC95 models
  - The estimated σ_u^2 for Kumb90 models
- The distribution of the idiosyncratic noise term. This is taken as:
  - The log of estimated σ_v^2 for BC95 models
  - The estimated σ_v^2 for Kumb90 models
- The starting values for the time-varying efficiency component. A vector of zeroes of the appropriate length is taken i.e. starting at no adjustment/trend for Kumb90 models and half-normal for all observations for BC95 models. This approach aligns the starting values of the efficiency time trend with the time invariant half-normal restriction of the model estimated in Step 2.

The Cobb-Douglas specification is then estimated.

Step 4 – estimate the translog time-varying efficiency SFA model

Having estimated the Cobb-Douglas specification in step 3, the translog specification is then estimated by using the resulting estimates (cost function, inefficiency distribution, idiosyncratic noise term distribution and time-varying efficiency component estimates), but extending the cost function to be a translog cost function (including the additional six cross/square terms) and inserting a vector of six zeros in the starting values for the cost function.

This results in the estimation starting at a good model fit (equal to that of the Cobb-Douglas model fit) ahead of allowing the model fit to improve by exploiting the flexibility of the translog specification. Crucially, the SFA components (inefficiency distribution, idiosyncratic noise distribution and time-varying efficiency component) are all set to reasonable values: these are the terms that can give rise to convergence issues if inappropriate starting values are used.

We also note the following:

- The BC95 starting values for the distribution of v and u should be in the form $\log(\sigma_v^2)$ and $\log(\sigma_u^2)$ as these are the terms used to represent the distribution of v and u for the BC95 model in the *sfpanel* package. The Kumb90 starting values for the distribution of v and u should be in the form σ_v^2 and σ_u^2 as these are the terms used to represent the distribution of v and u for the Kumb90 model in the *sfpanel* package.
- Though not applied for the models estimated, when estimating the heteroscedasticity models we would set the starting value of the impact of size on idiosyncratic error variance to zero.
- In general, we find it useful to use the command "matrix list e(b)" after estimating a model – this provides information on the actual parameter estimates used in the model, and therefore how the starting values should be constructed and inputted.

B Approach to estimating efficiency in BC95 models

This Appendix sets out an approach to estimating efficiency in BC95 models that avoids an error that can be encountered in certain BC95 models, in particular the AJTT models.

The *sfpanel* package contains a built-in post-estimation command to estimate efficiency of firms.[90] However, that approach may fail to produce efficiency estimates for some firms for some years when using the BC95 model, in particular the AJTT specifications of the BC95 model.

Inspection of the sfpanel-p.ado file provides the approach used by the *sfpanel* package to generate estimates of the efficiency using the "predict efficiency, bc" command. We note that Quantonomics used the same command in estimating efficiency.

This command applies the Battese-Coelli (1988) approach to estimating efficiency. As set out in the ado file,[91] the generated efficiency is calculated as:

$$\frac{normal\left(-\sigma_1 + \frac{\mu_1}{\sigma_1}\right)}{normal\left(\frac{\mu_1}{\sigma_1}\right)} \times \exp\left(-\mu_1 + \frac{\sigma_1^2}{2}\right),$$

where

$$\sigma_1^2 = \frac{\sigma_u^2 \sigma_v^2}{\sigma_u^2 + \sigma_v^2},$$

and

$$\mu_1 = \frac{Resid_{it}\sigma_u^2 + \mu_{it}\sigma_v^2}{\sigma_u^2 + \sigma_v^2}$$

The residual, $Resid_{it}$, is calculated as the difference between the estimated cost function and the actual cost for the observation (rather than the difference between the actual cost and the fitted cost incorporating the estimated efficiency, as presented in residual plots throughout the Quantonomics report and this report).

The mean of the non-truncated inefficiency distribution, μ_{it}, is calculated as per the relevant BC95 specification, in general varying by DNSP and time.

We note that the $normal(\cdot)$ function is defined as the cumulative distribution function of the normal distribution. For large negative values, i.e. for large negative $\frac{\mu_1}{\sigma_1}$, this would be close to zero. For sufficiently negative values, the value of the normal distribution is saved in Stata as precisely zero.[92] This is due to a limitation on how Stata stores numbers during calculations. This results in an undefined efficiency as the truncated normal distribution is unable to be calculated by Stata.

To circumvent this we rewrite the equation above as:

$$\exp\left(lnnormal\left(-\sigma_1 + \frac{\mu_1}{\sigma_1}\right) - lnnormal\left(\frac{\mu_1}{\sigma_1}\right)\right) \times \exp\left(-\mu_1 + \frac{\sigma_1^2}{2}\right)$$

90 This is performed using the command "predict efficiency, bc".

91 See lines 600-605, with various terms defined lines 448, 450 and 527.

92 When using double precision numbers, as used for calculations in Stata, $\frac{\mu_1}{\sigma_1}$ below -37.6 results in a cumulative distribution value of zero.

By using the *lnnormal* function in Stata we avoid the issue whereby the value of the distribution is rounded to zero. Calculating the efficiency using the approach yields precisely the same efficiencies in circumstances where the built-in approach generated efficiencies, and provides efficiencies for DNSP year combinations that failed to produce valid efficiency estimates.

To demonstrate that this approach is straightforward to implement, we provide sample Stata code below for the BC95 models:[93]

Stata code to estimate the efficiency scores for BC95 models

```
predict double lvchat, xb
gen double residual=lvc-lvchat
predict double mu, eq(Mu) xb
gen double sigma-v2= e(sigma-v)^2
gen double sigma-u2= e(sigma-u)^2
gen double mu1 = (residual*sigma-u2+mu*sigma-v2)/(sigma-u2+sigma-v2)
gen double sigma1 = (sqrt(sigma-u2)*sqrt(sigma-v2))/(sqrt(sigma-u2+sigma-v2))
gen double efficiency= exp(-mu1+0.5*sigma1^2)*exp(lnnormal(-sigma1+mu1/sigma1)-
lnnormal(mu1/sigma1))
```

Source: Frontier Economics

93 Note that for heteroscedasticity models an alternative expression for sigma-v2 would be required. This is trivial to implement.

C Additional model results

Efficiency patterns

In this section we provide figures detailing the estimated time-varying efficiency of Australian DNSPs under the various models considered.

Figure 16: DNSP efficiency scores over time for BC95-JTT-HN-TL

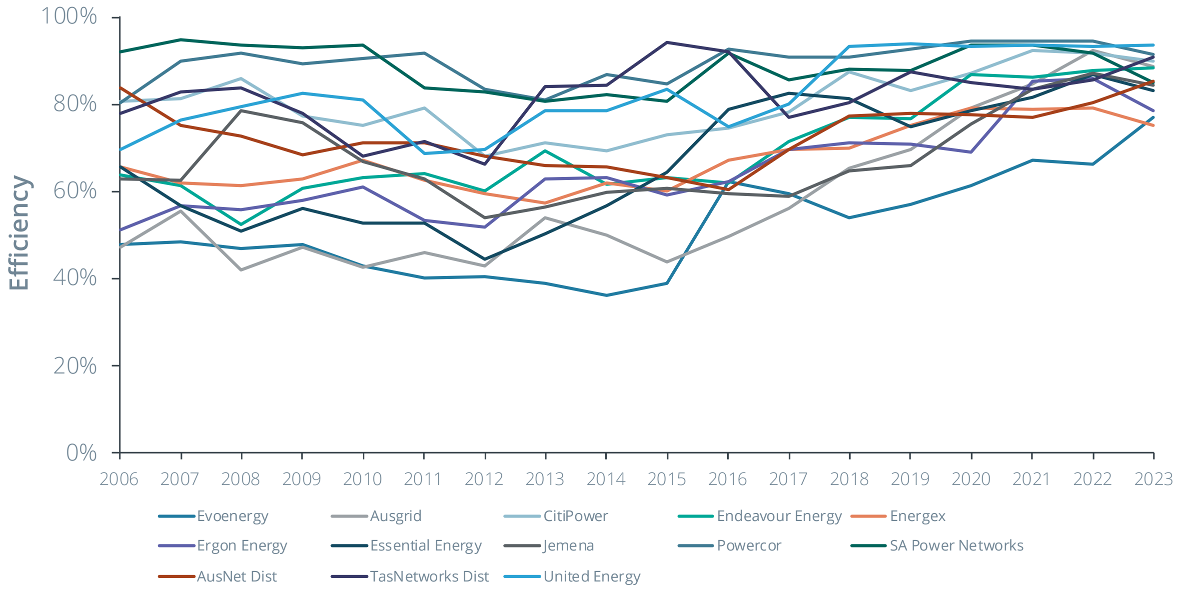
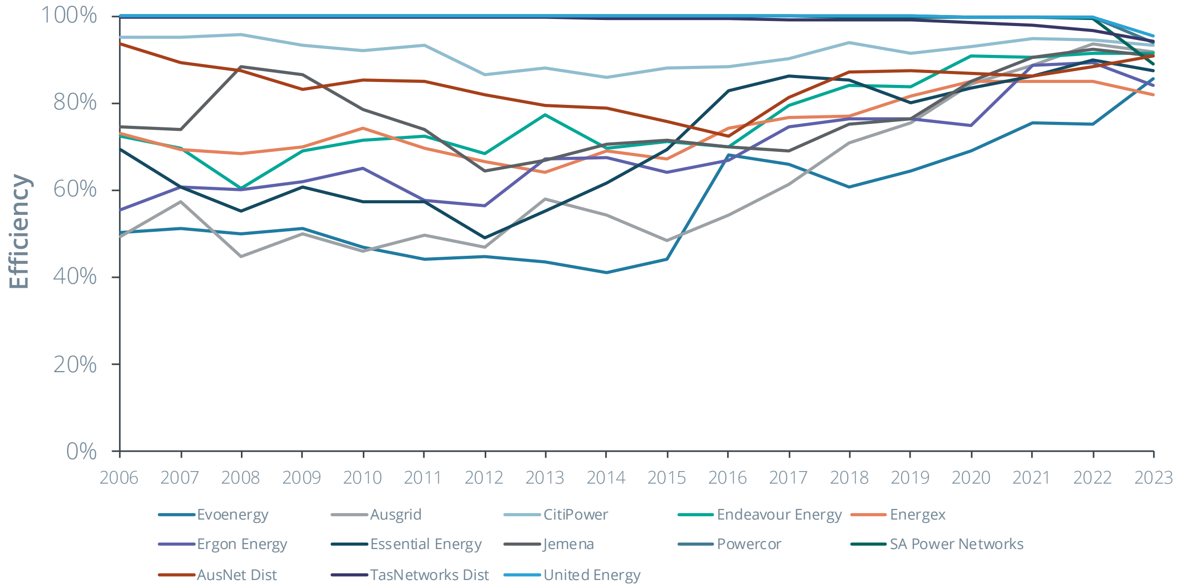
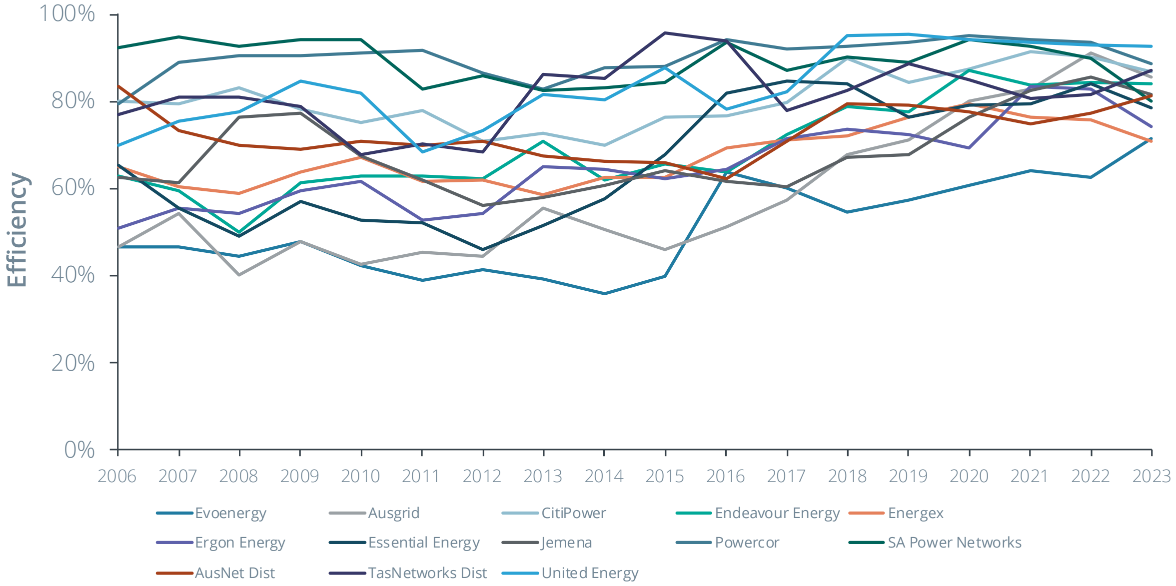


Figure 17: DNSP efficiency scores over time for BC95-AJTT-HN-TL



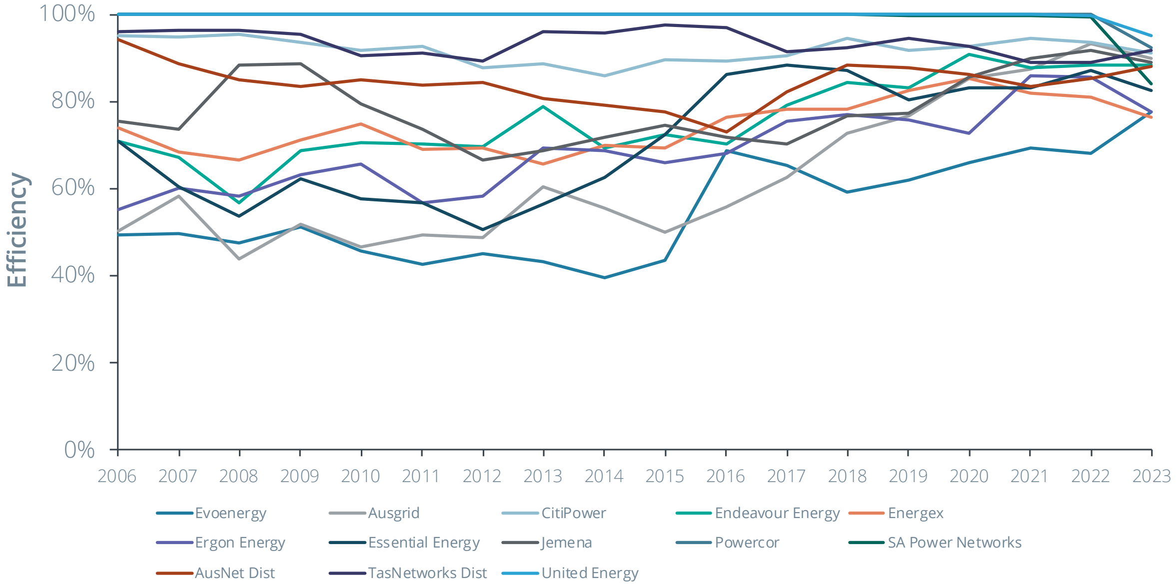
Source: Frontier Economics.

Figure 18: DNSP efficiency scores over time for BC95-JTT-HN-GTC-TL



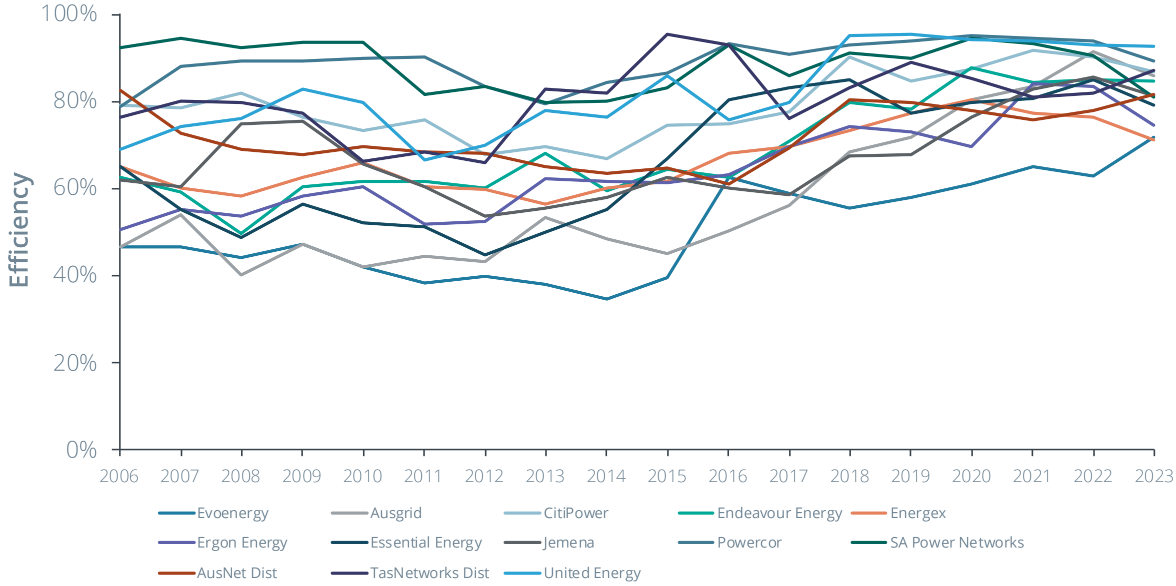
Source: Frontier Economics.

Figure 19: DNSP efficiency scores over time for BC95-AJTT-HN-GTC-TL



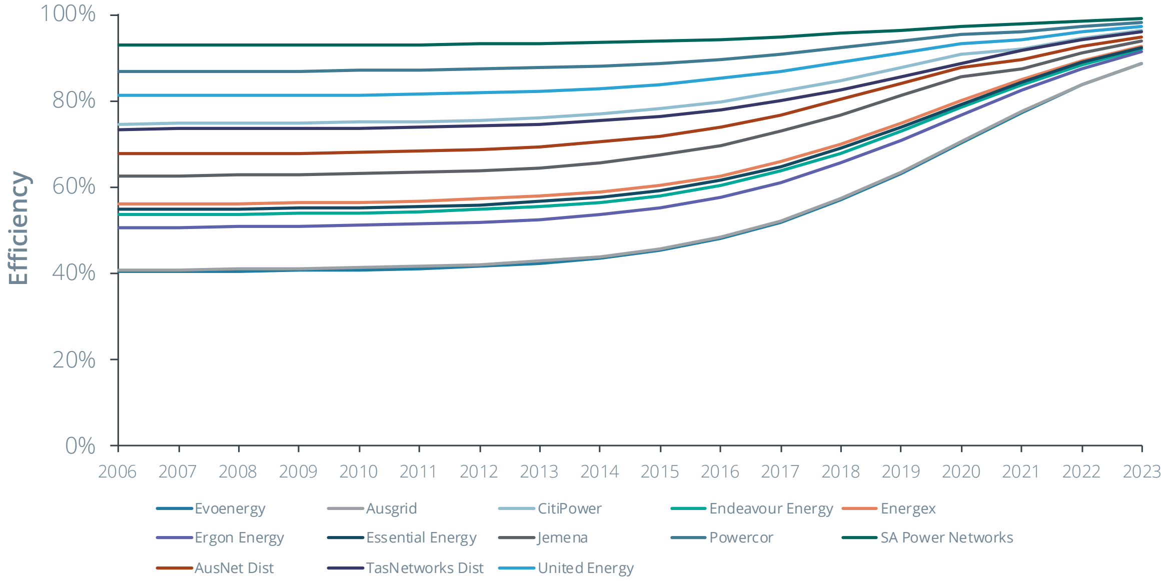
Source: Frontier Economics.

Figure 20: DNSP efficiency scores over time for BC95-AGEC-HN-GTC-TL



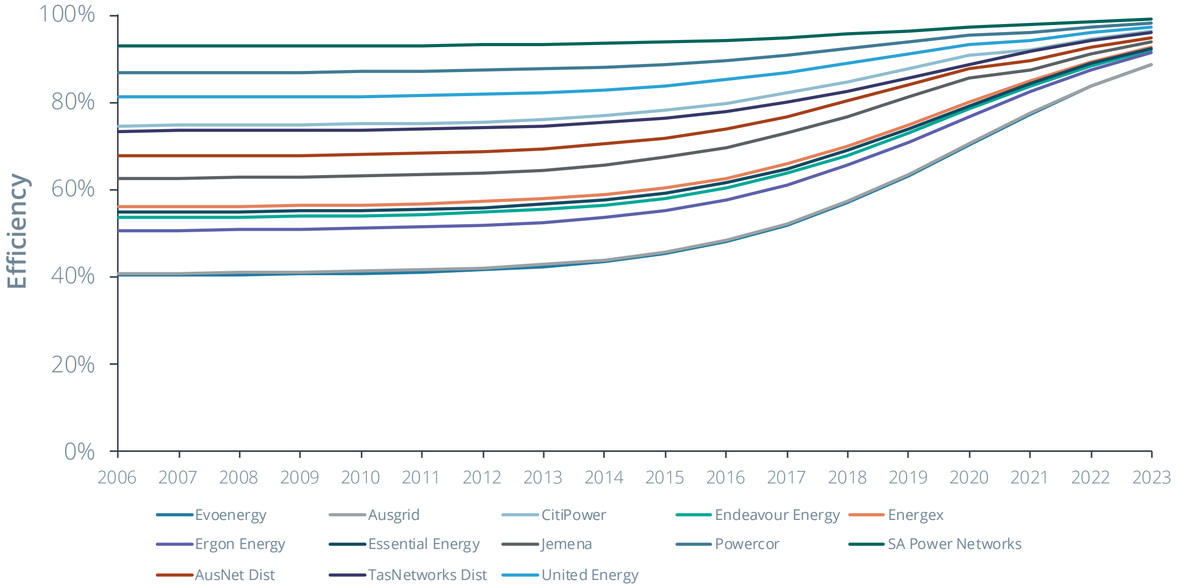
Source: Frontier Economics.

Figure 21: DNSP efficiency scores over time for Kumb-JTT-HN-TL



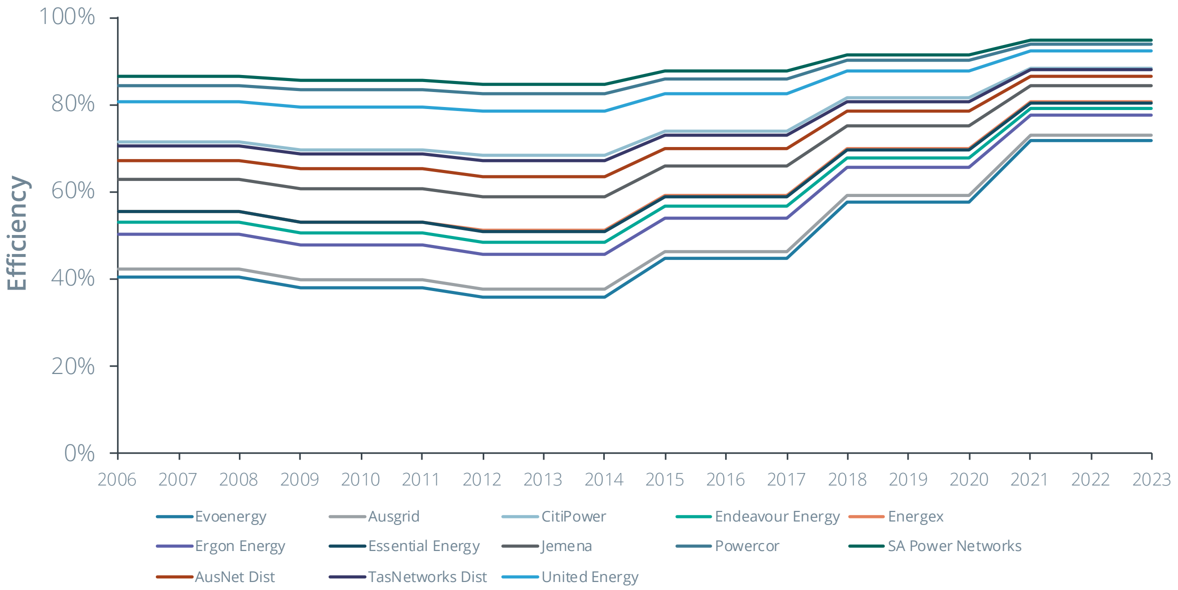
Source: Frontier Economics.

Figure 22: DNSP efficiency scores over time for Kumb-AGEC-HN-TL



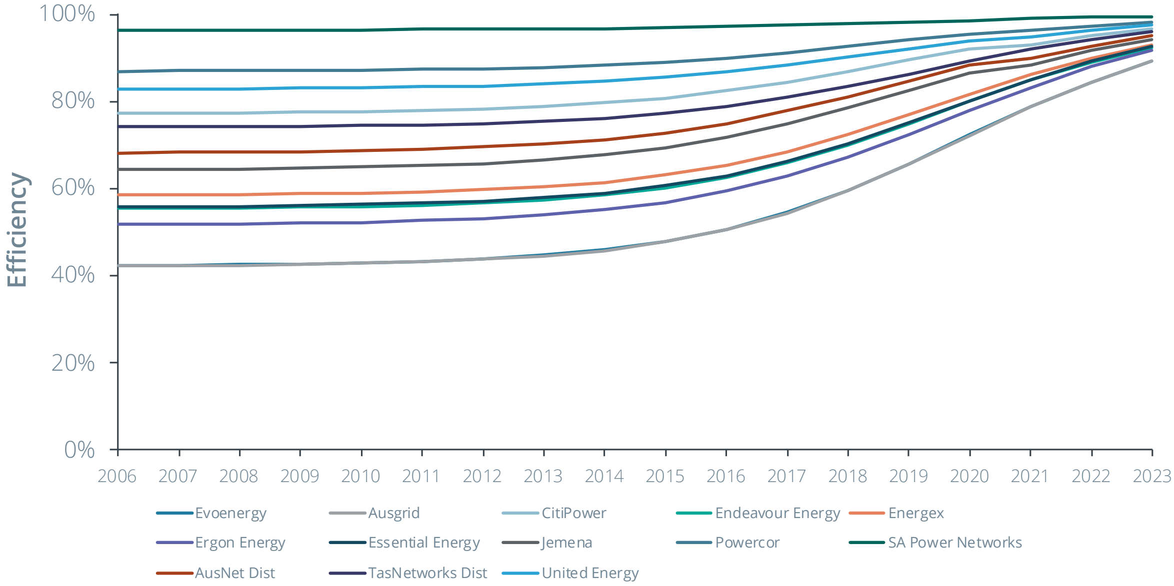
Source: Frontier Economics.

Figure 23: DNSP efficiency scores over time for Kumb-AGECJUR-HN-TL



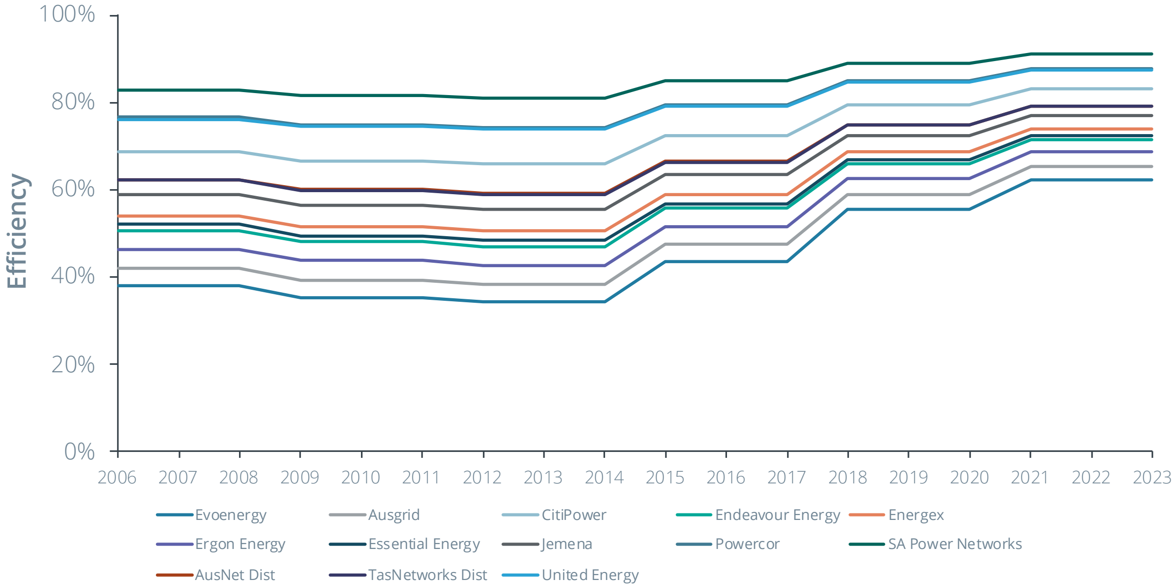
Source: Frontier Economics.

Figure 24: DNSP efficiency scores over time for Kumb-JTT-HN-GTC-TL



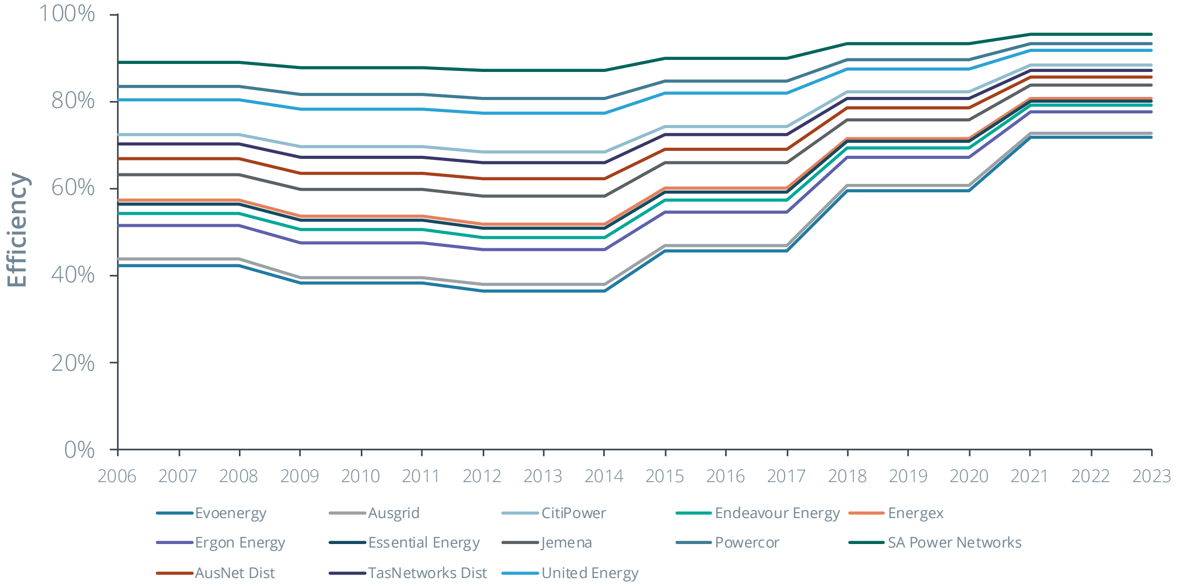
Source: Frontier Economics.

Figure 25: DNSP efficiency scores over time for Kumb-AGEC-HN-GTC-TL



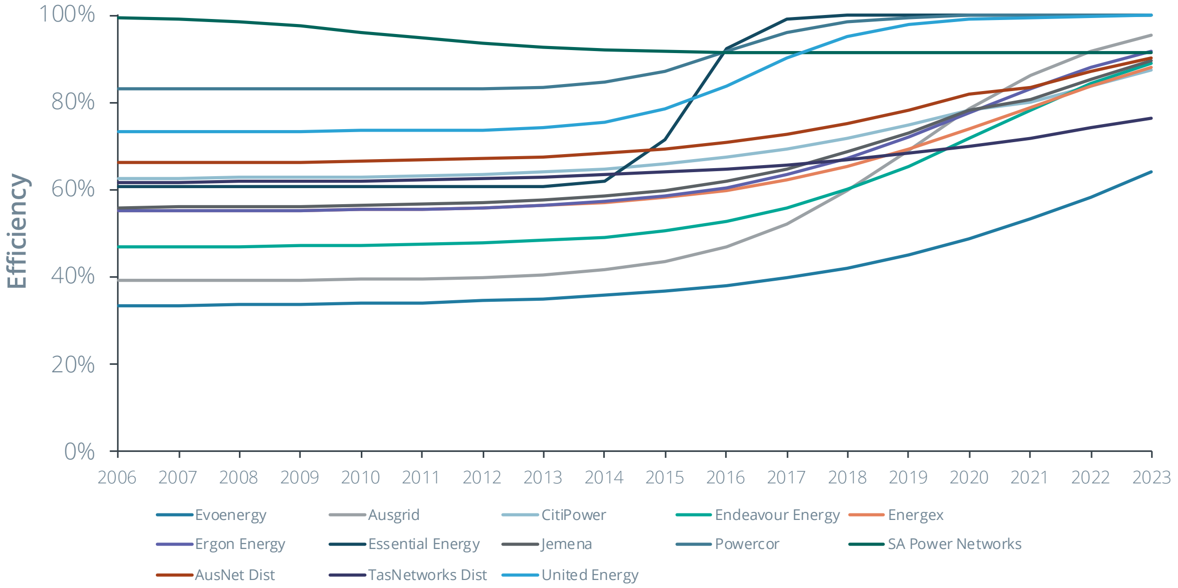
Source: Frontier Economics.

Figure 26: DNSP efficiency scores over time for Kumb-AGECJUR-HN-GTC-TL



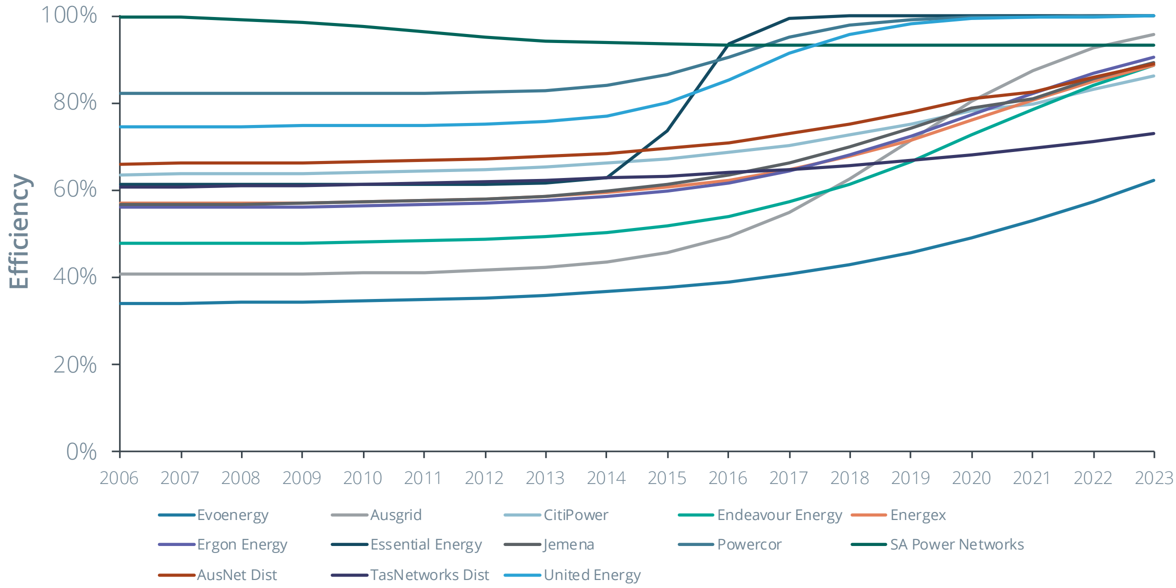
Source: Frontier Economics.

Figure 27: DNSP efficiency scores over time for Kumb-AJTT-HN-TL



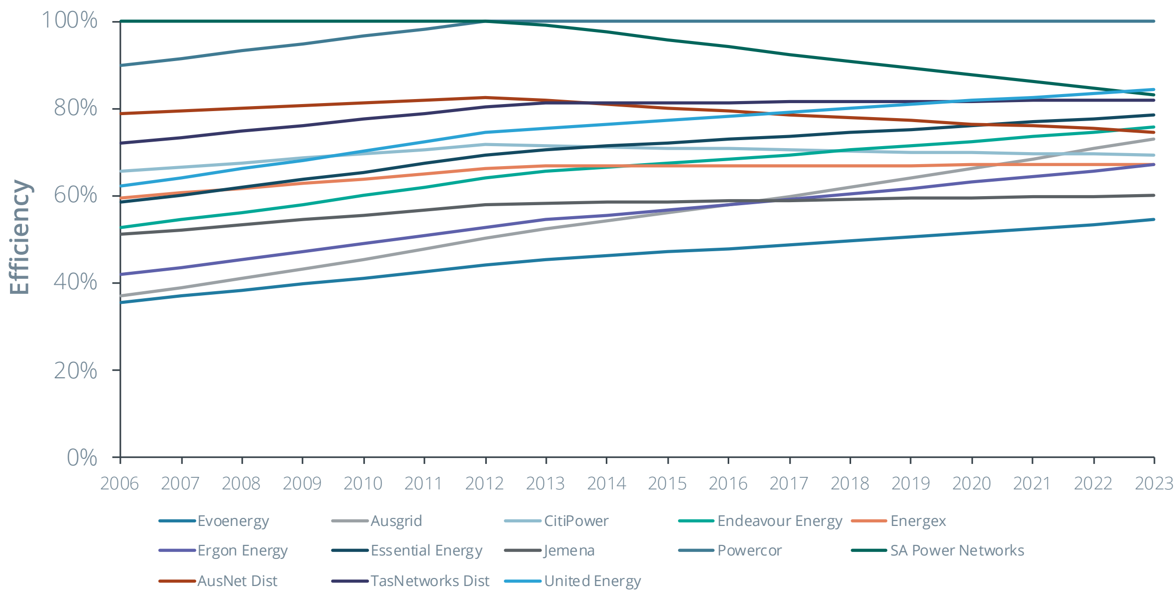
Source: Frontier Economics.

Figure 28: DNSP efficiency scores over time for Kumb-AJTT-HN-GTC-TL



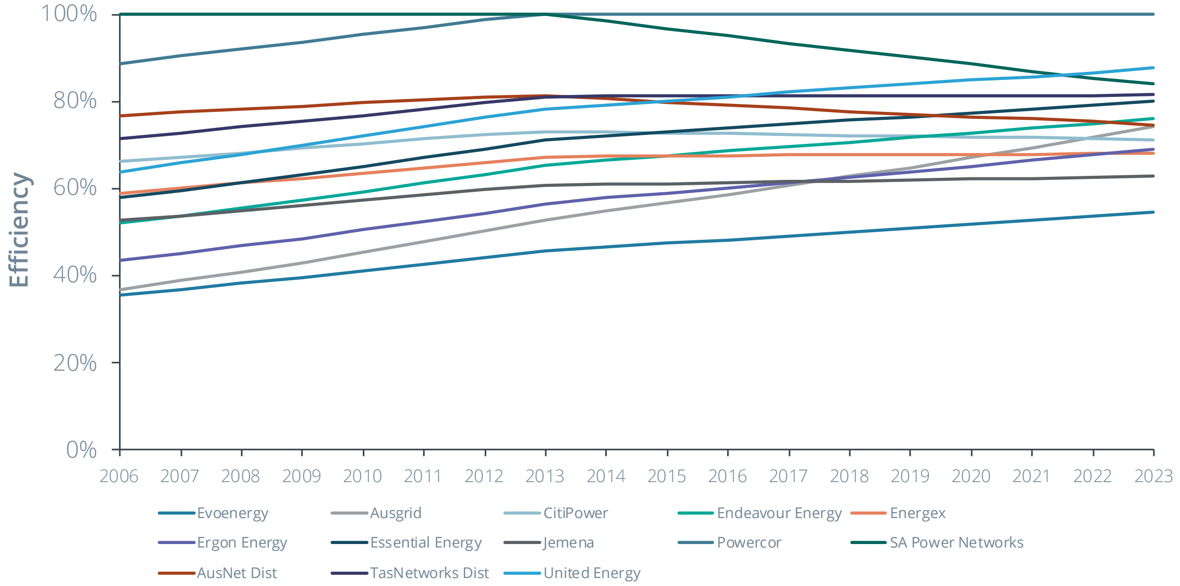
Source: Frontier Economics.

Figure 29: DNSP efficiency scores over time for LSE-ADTT-TL



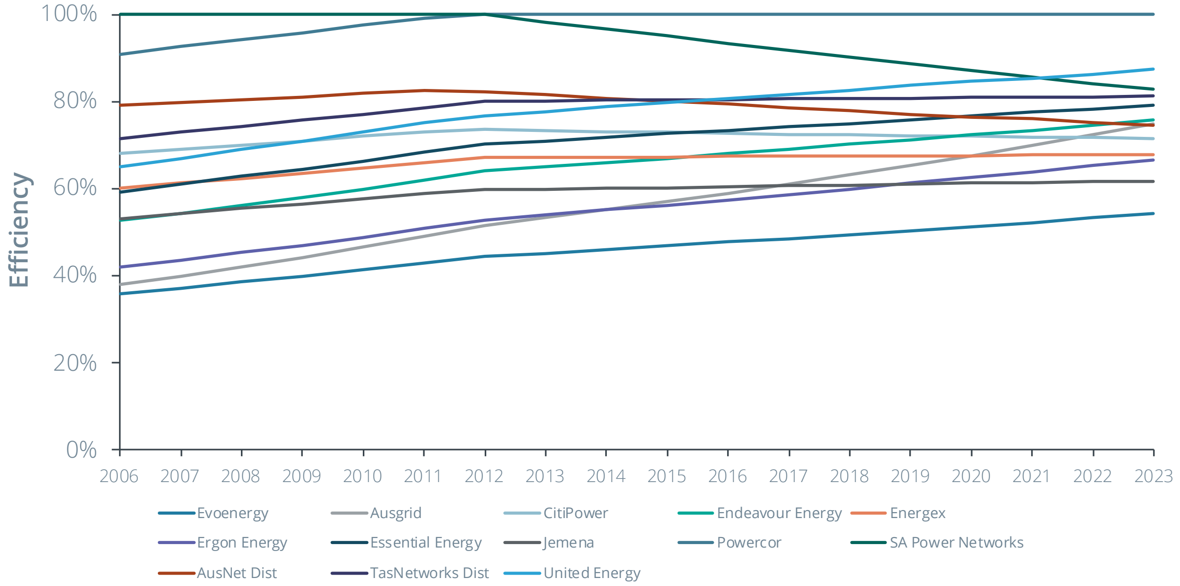
Source: Frontier Economics.

Figure 30: DNSP efficiency scores over time for LSE-ADTT-GTC-TL



Source: Frontier Economics.

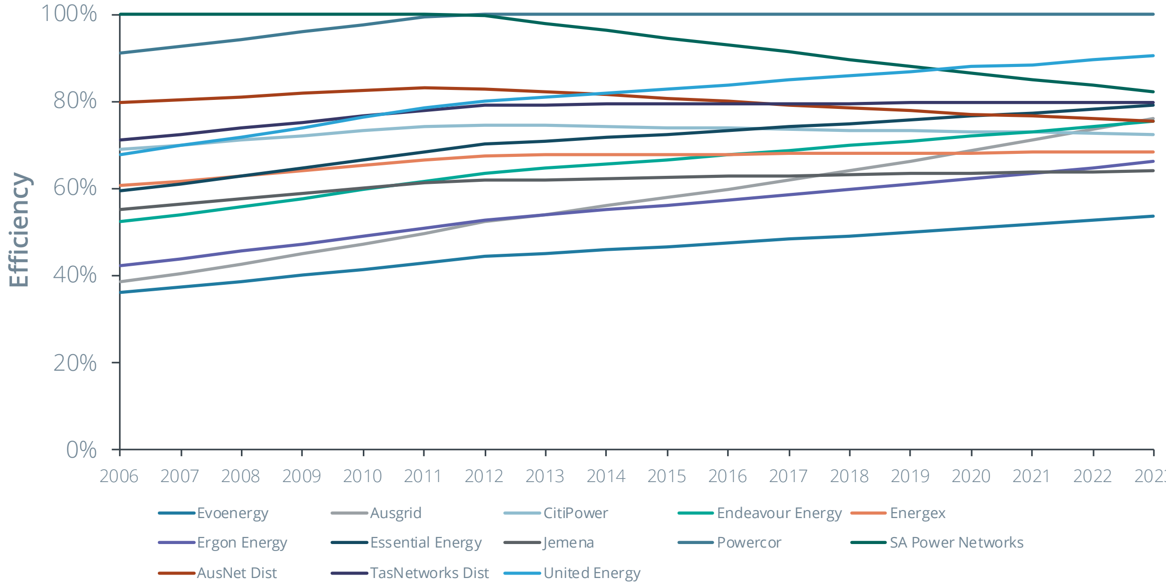
Figure 31: DNSP efficiency scores over time for LSE-AJTT-NZ-TL



Source: Frontier Economics.

Note: The residuals and coefficients estimate for the New Zealand and Ontario versions of this model are identical, with exception only being the overall year trend. This raises questions of how the roll-forward model should be applied, if the time trend varies depending on the country chosen.

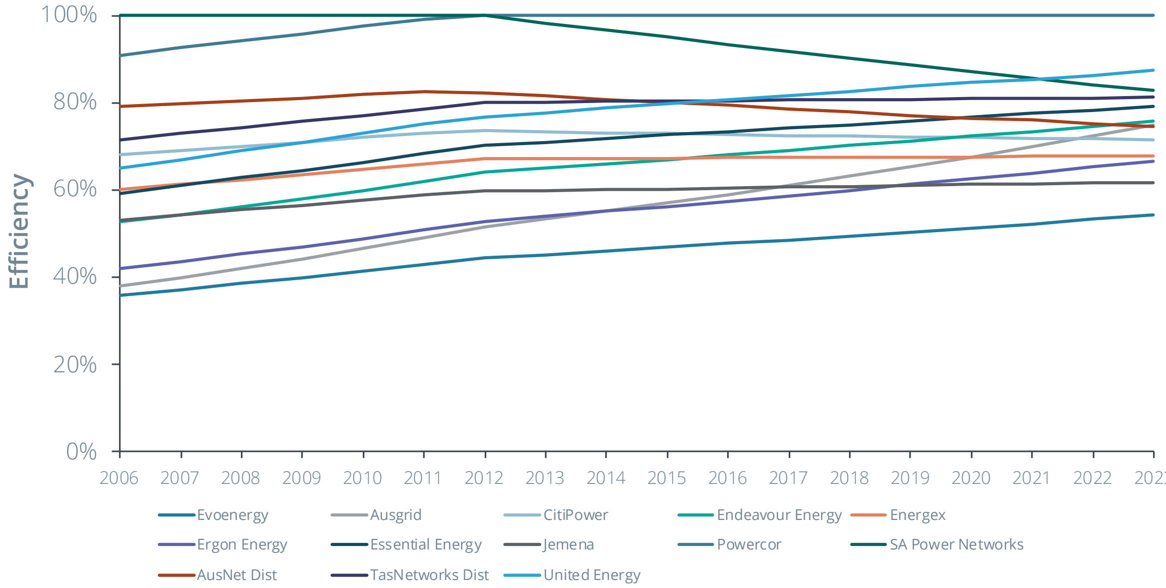
Figure 32: DNSP efficiency scores over time for LSE-AJTT-NZ-GTC-TL



Source: Frontier Economics.

Note: The residuals and coefficients estimate for the New Zealand and Ontario versions of this model are identical, with exception only being the overall year trend. This raises questions of how the roll-forward model should be applied, if the time trend varies depending on the country chosen.

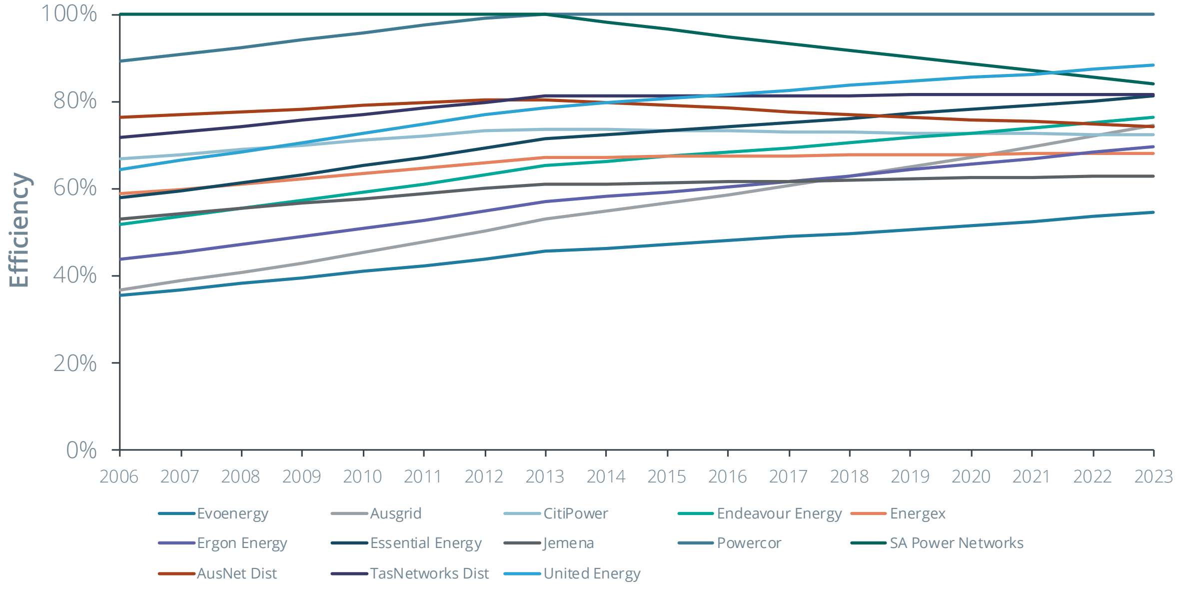
Figure 33: DNSP efficiency scores over time for LSE-AJTT-ONT-TL



Source: Frontier Economics.

Note: The residuals and coefficients estimate for the New Zealand and Ontario versions of this model are identical, with exception only being the overall year trend. This raises questions of how the roll-forward model should be applied, if the time trend varies depending on the country chosen.

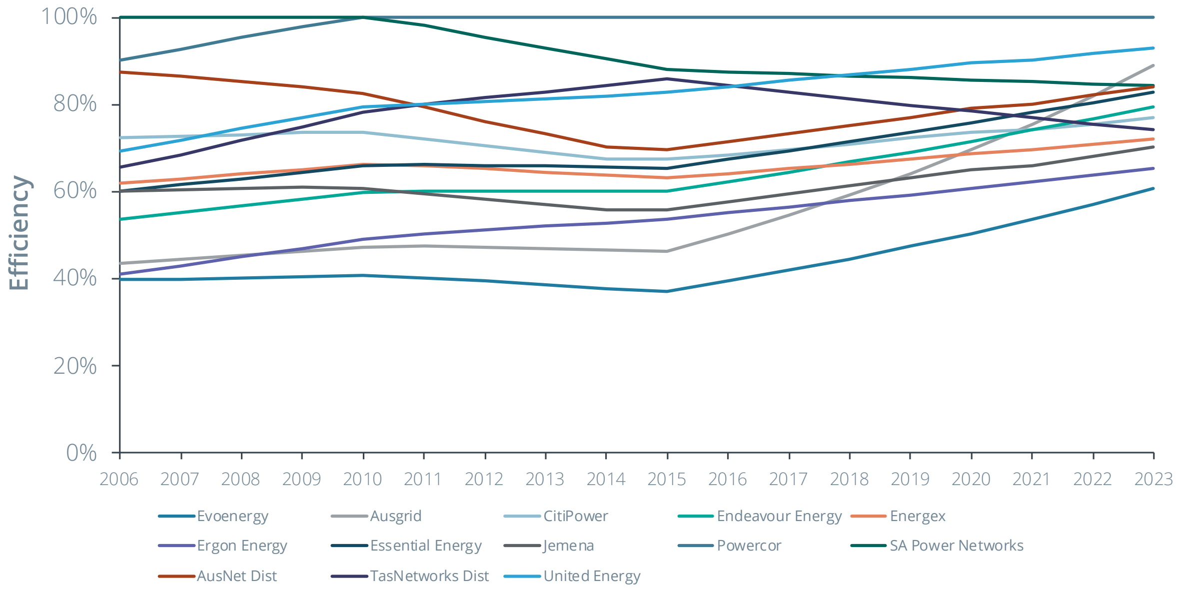
Figure 34: DNSP efficiency scores over time for LSE-AJTT-ONT-GTC-TL



Source: Frontier Economics.

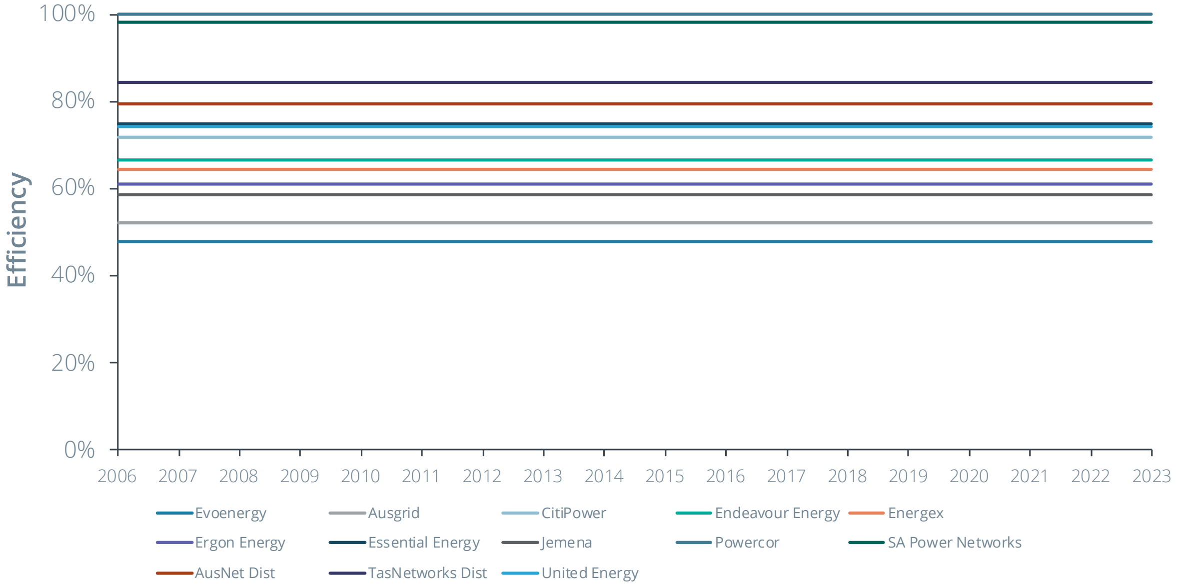
Note: The residuals and coefficients estimate for the New Zealand and Ontario versions of this model are identical, with exception only being the overall year trend. This raises questions of how the roll-forward model should be applied, if the time trend varies depending on the country chosen.

Figure 35: DNSP efficiency scores over time for LSE-AJTTBREAK-NZ-TL



Source: Frontier Economics.

Figure 36: DNSP efficiency scores over time for LSE-TI-TL



Source: Frontier Economics.

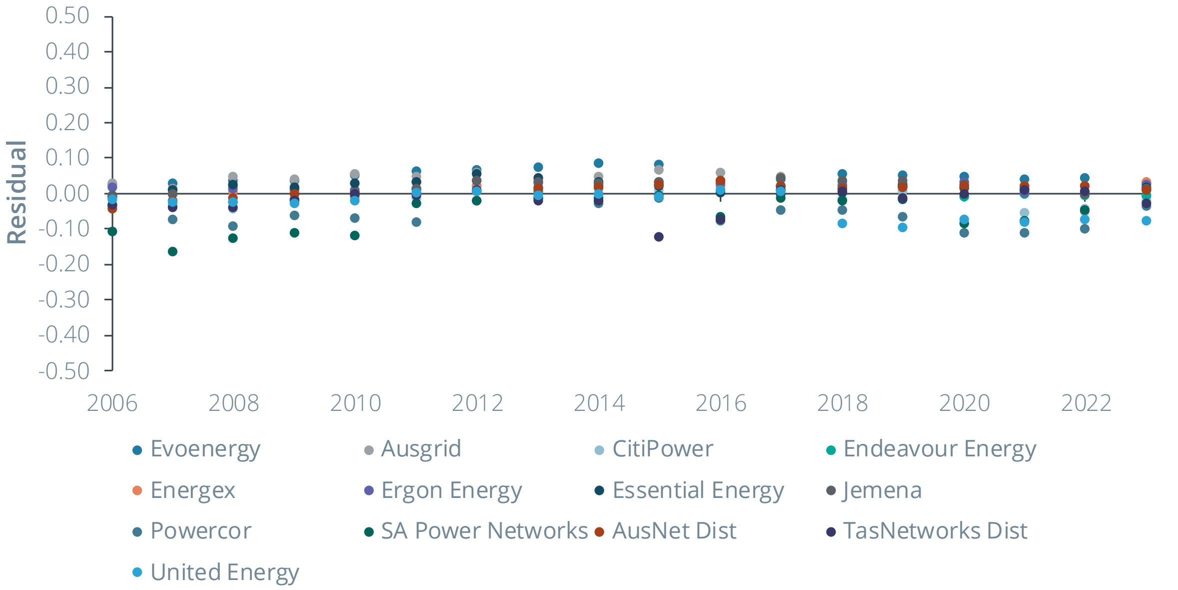
Figure 37: DNSP efficiency scores over time for SFA-TI-TL

Source: Frontier Economics.

Residual patterns

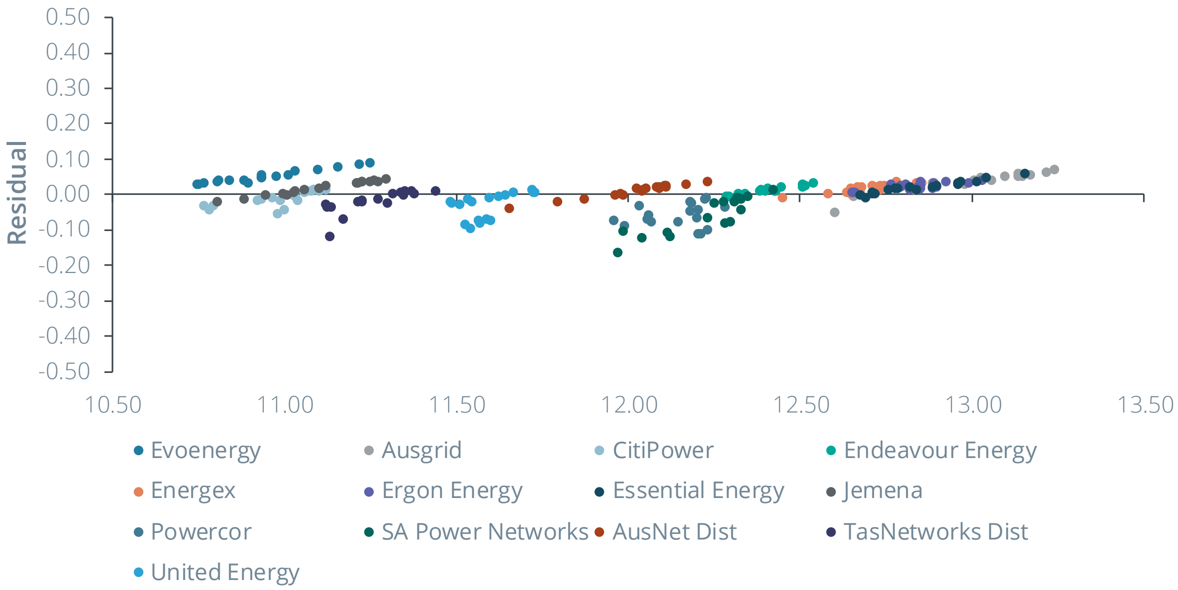
In this section we provide figures detailing the residuals generated under the various models considered.

Figure 38: Residuals plot over time for BC95-JTT-HN-TL



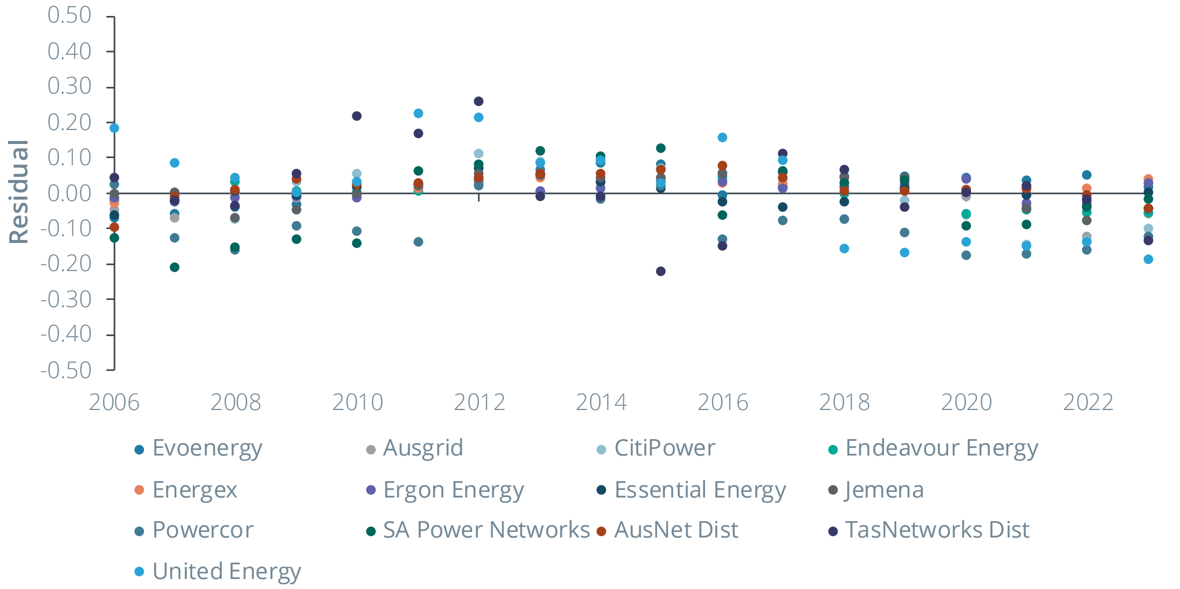
Source: Frontier Economics.

Figure 39: Residuals plot over fitted value for BC95-JTT-HN-TL



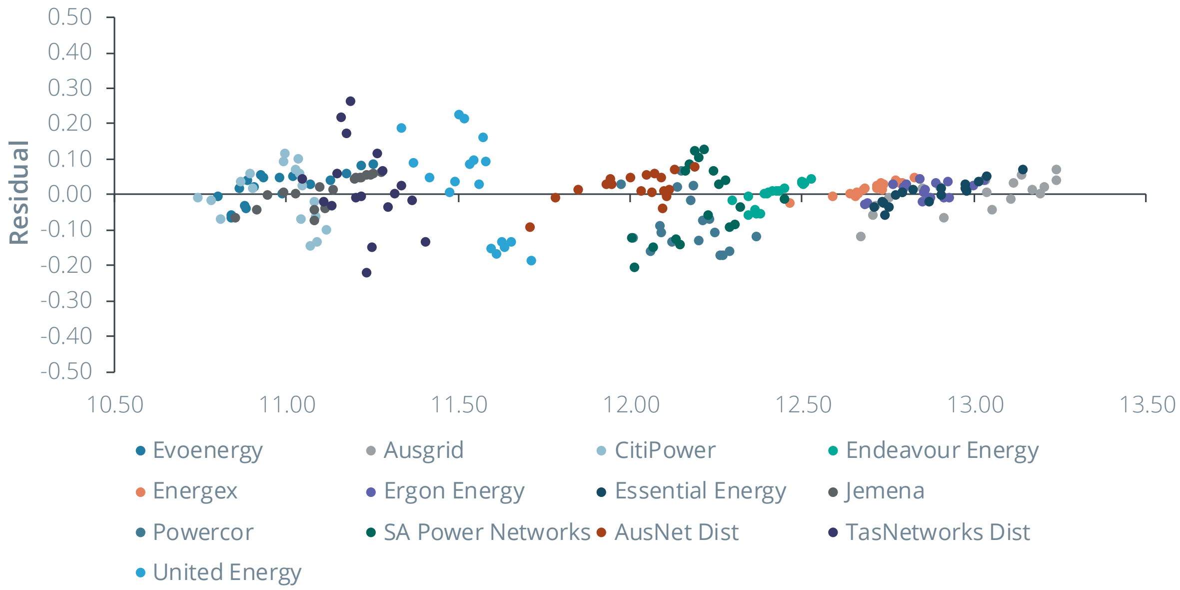
Source: Frontier Economics.

Figure 40: Residuals plot over time for BC95-AJTT-HN-TL



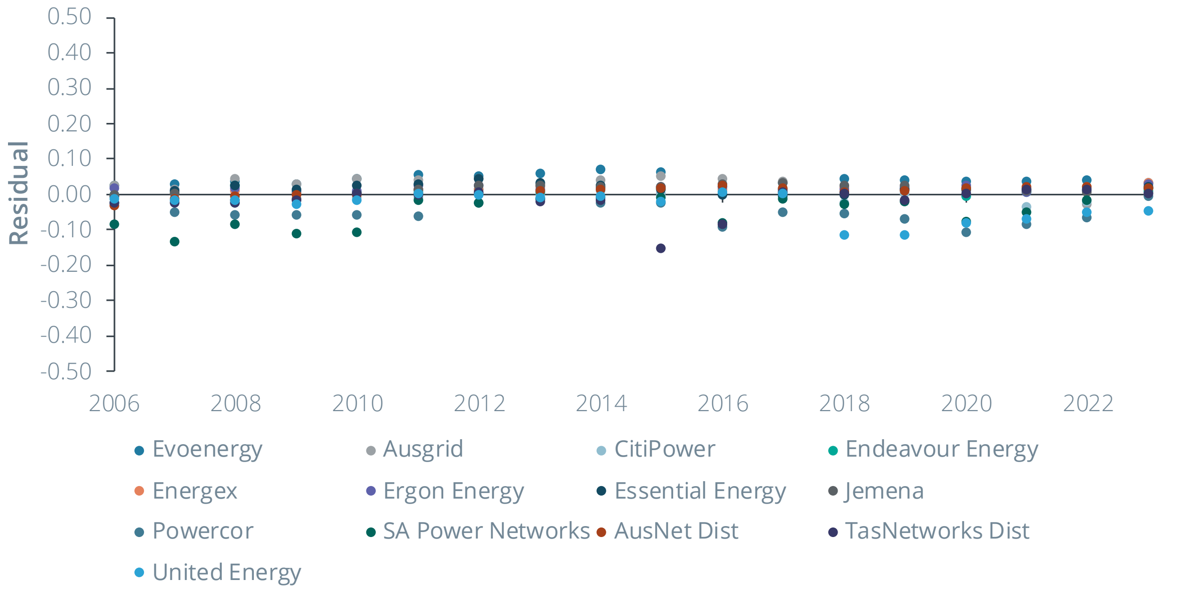
Source: Frontier Economics.

Figure 41: Residuals plot over fitted value for BC95-AJTT-HN-TL



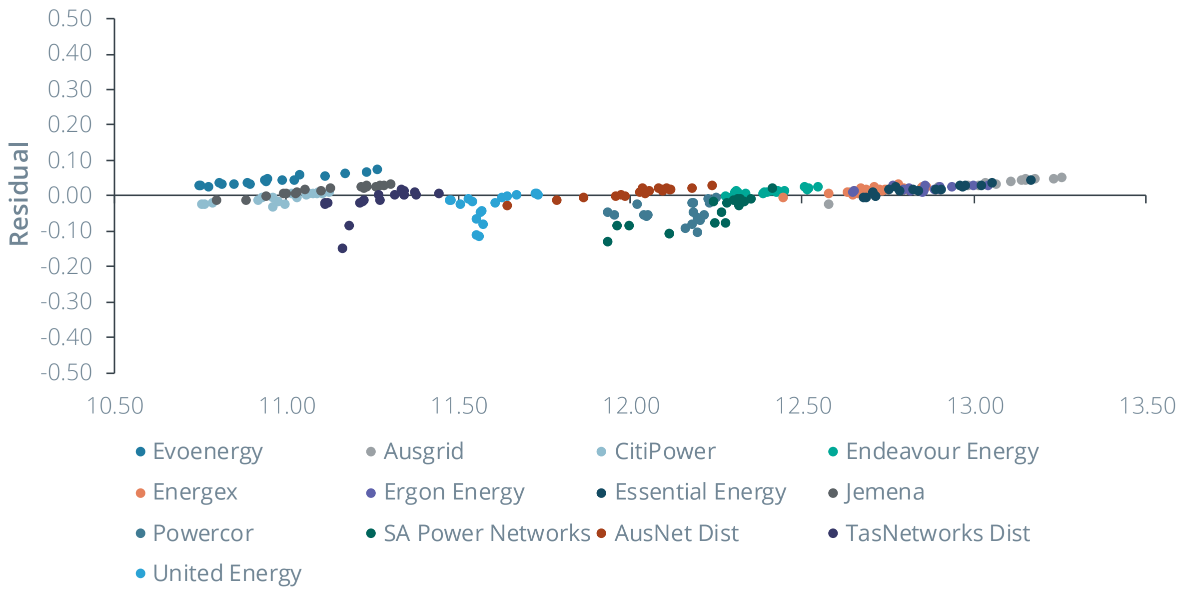
Source: Frontier Economics.

Figure 42: Residuals plot over time for BC95-JTT-HN-GTC-TL



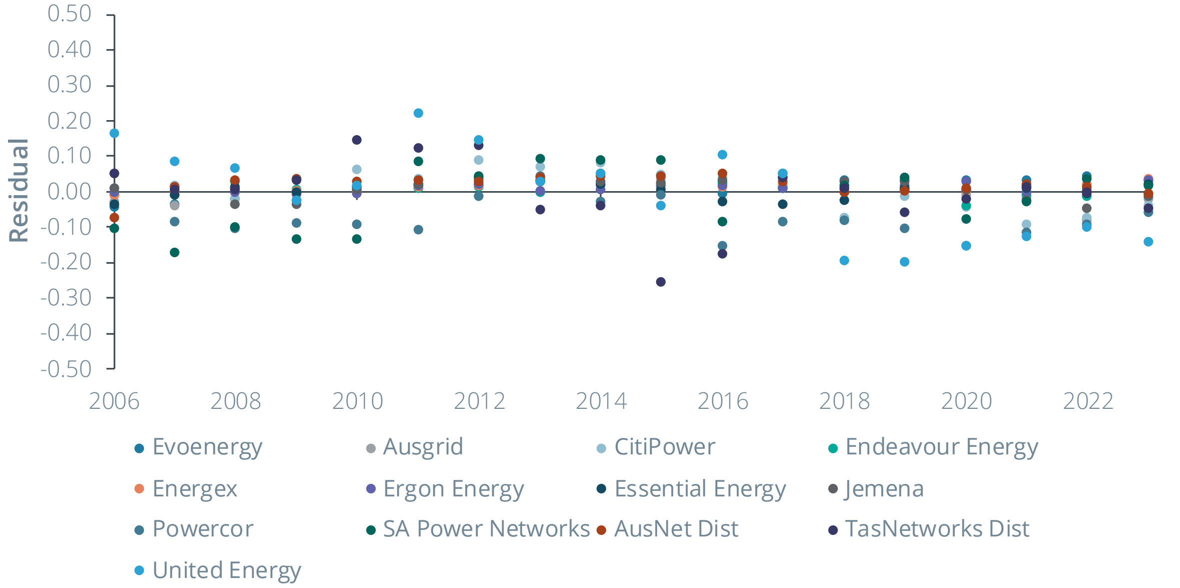
Source: Frontier Economics.

Figure 43: Residuals plot over fitted value for BC95-JTT-HN-GTC-TL



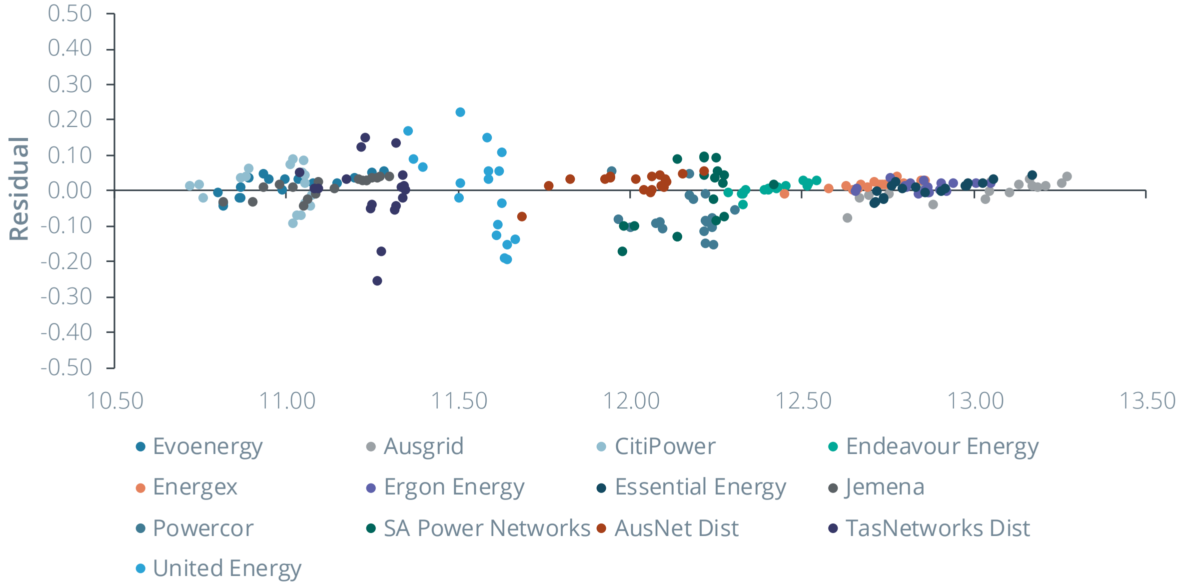
Source: Frontier Economics.

Figure 44: Residuals plot over time for BC95-AJTT-HN-GTC-TL



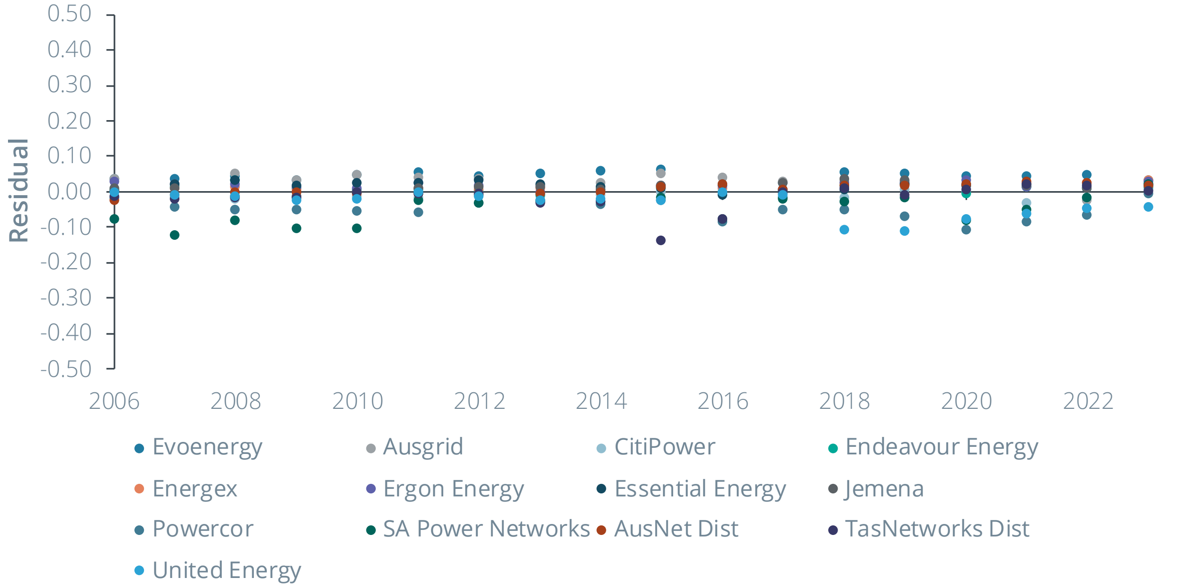
Source: Frontier Economics.

Figure 45: Residuals plot over fitted value for BC95-AJTT-HN-GTC-TL



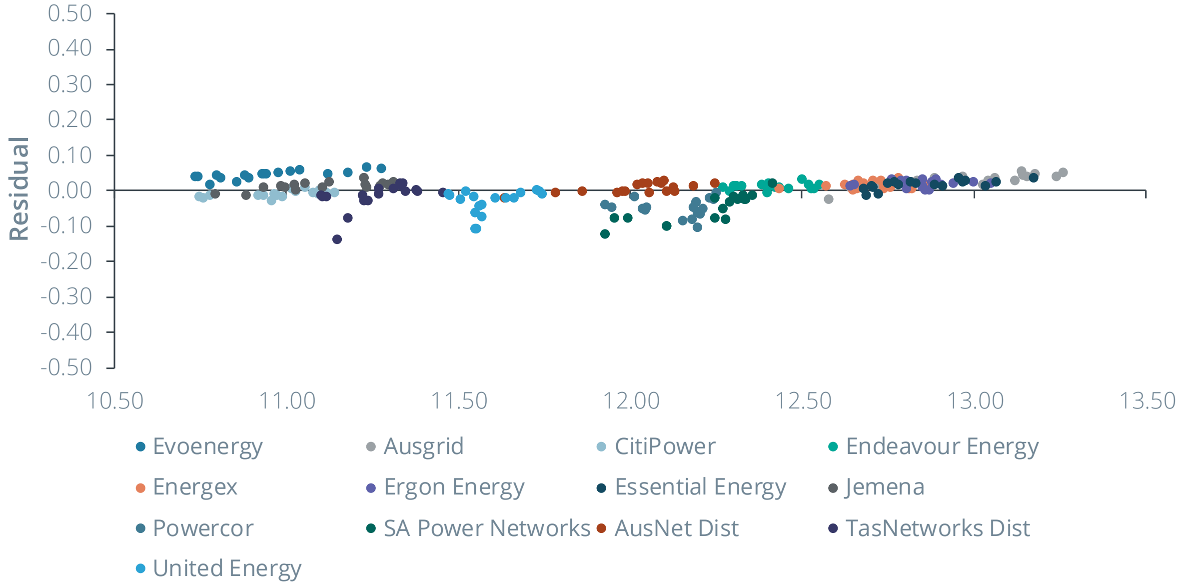
Source: Frontier Economics.

Figure 46: Residuals plot over time for BC95-AGEC-HN-GTC-TL



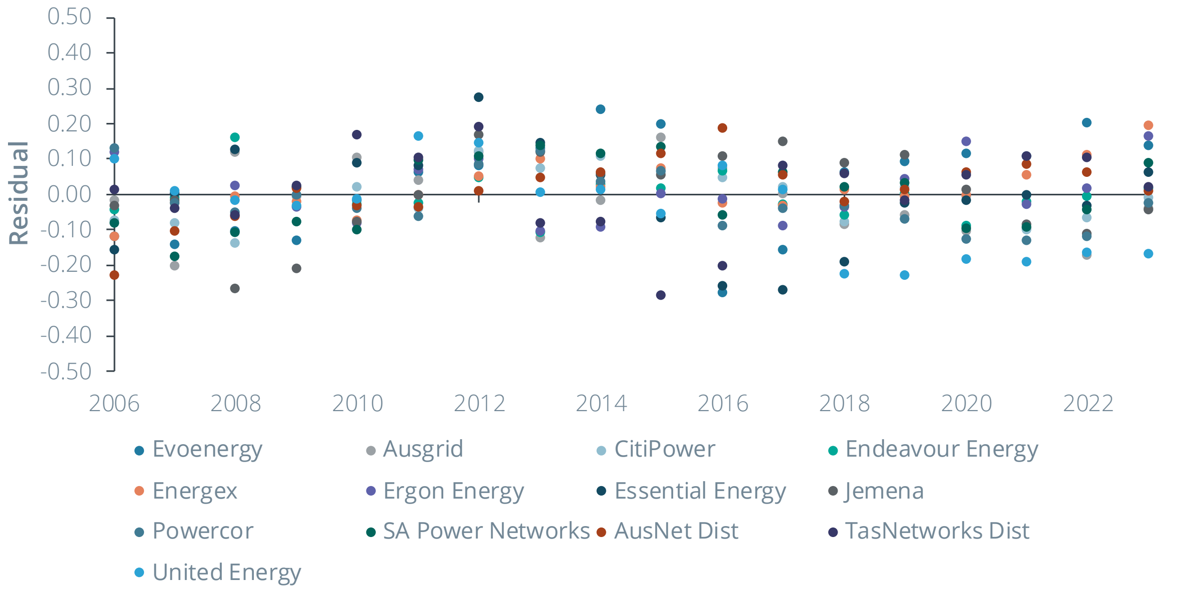
Source: Frontier Economics.

Figure 47: Residuals plot over fitted value for BC95-AGEC-HN-GTC-TL



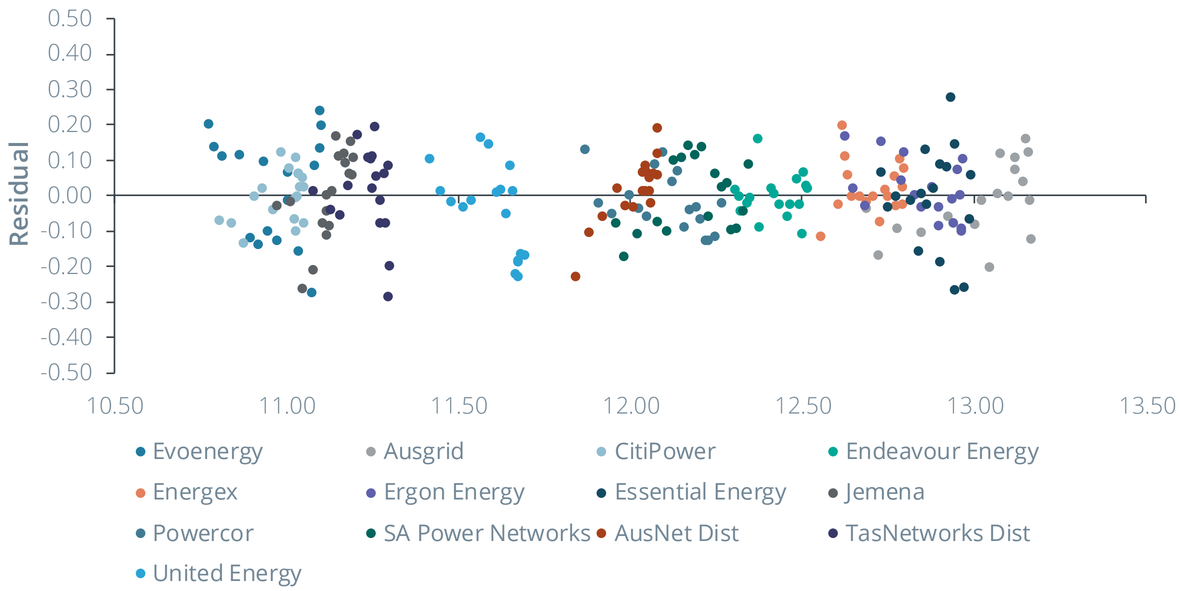
Source: Frontier Economics.

Figure 48: Residuals plot over time for Kumb-JTT-HN-TL



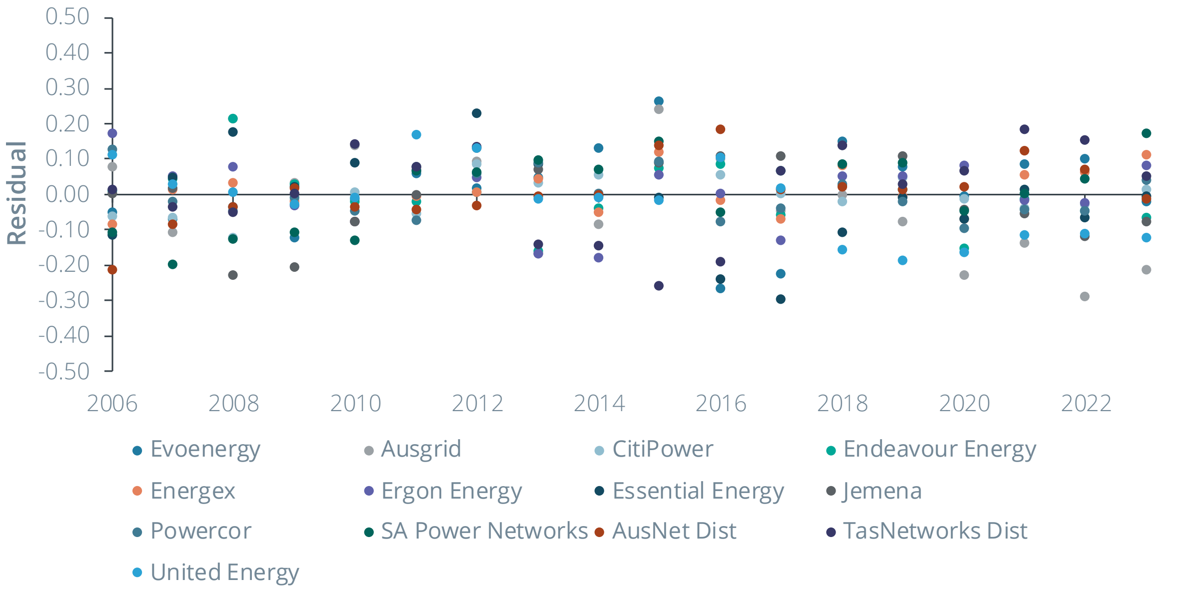
Source: Frontier Economics.

Figure 49: Residuals plot over fitted value for Kumb-JTT-HN-TL



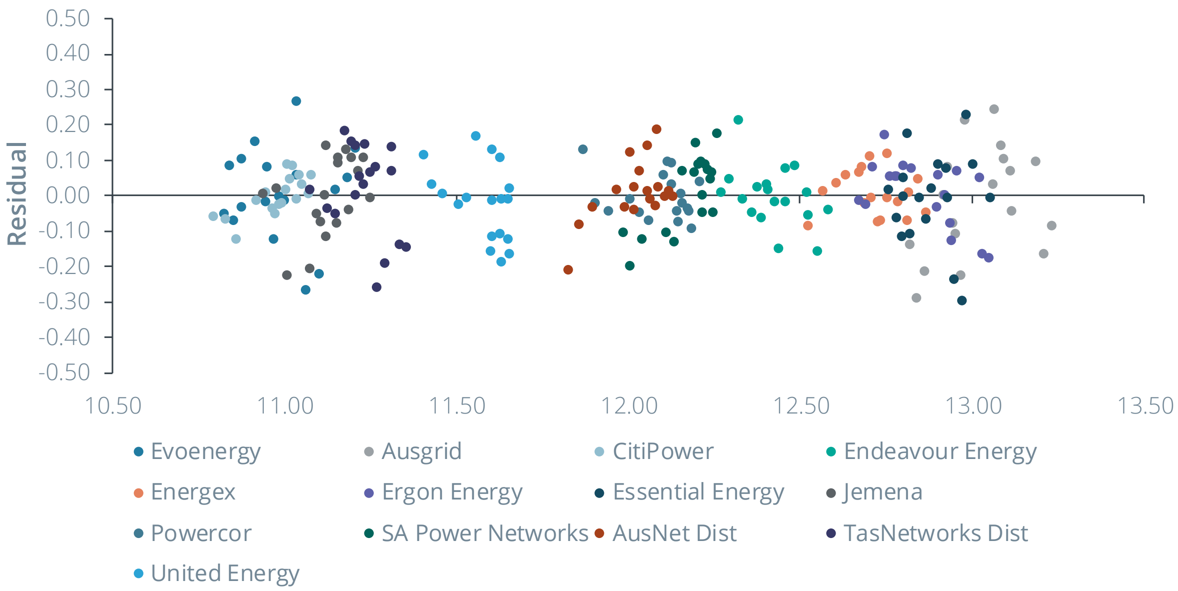
Source: Frontier Economics.

Figure 50: Residuals plot over time for Kumb-AGEC-HN-TL



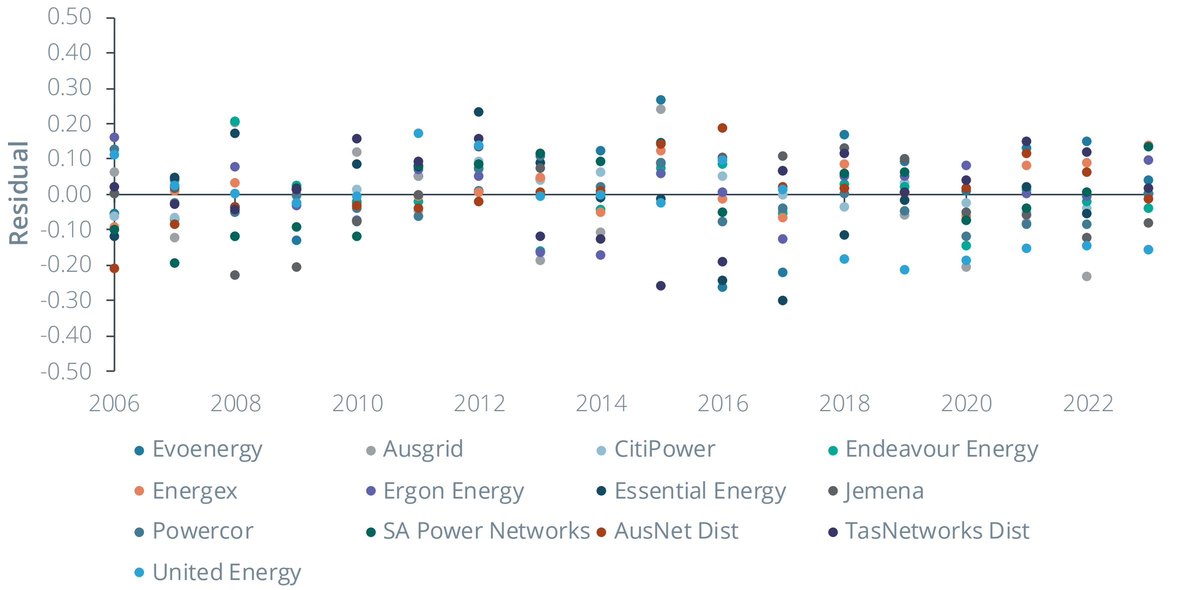
Source: Frontier Economics.

Figure 51: Residuals plot over fitted value for Kumb-AGEC-HN-TL



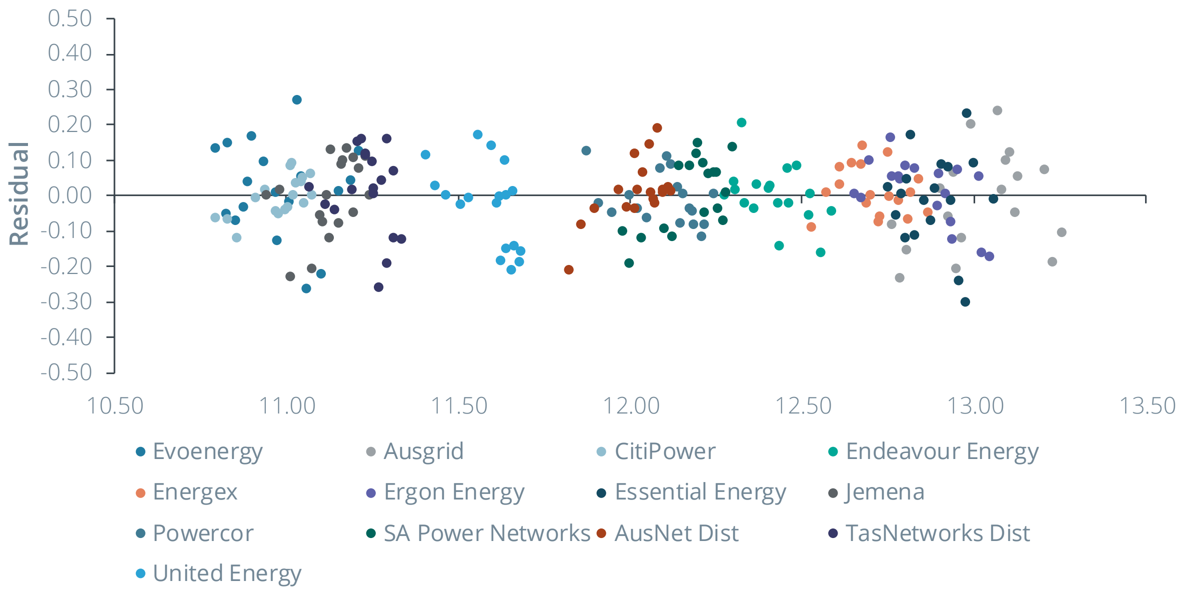
Source: Frontier Economics.

Figure 52: Residuals plot over time for Kumb-AGECJUR-HN-TL



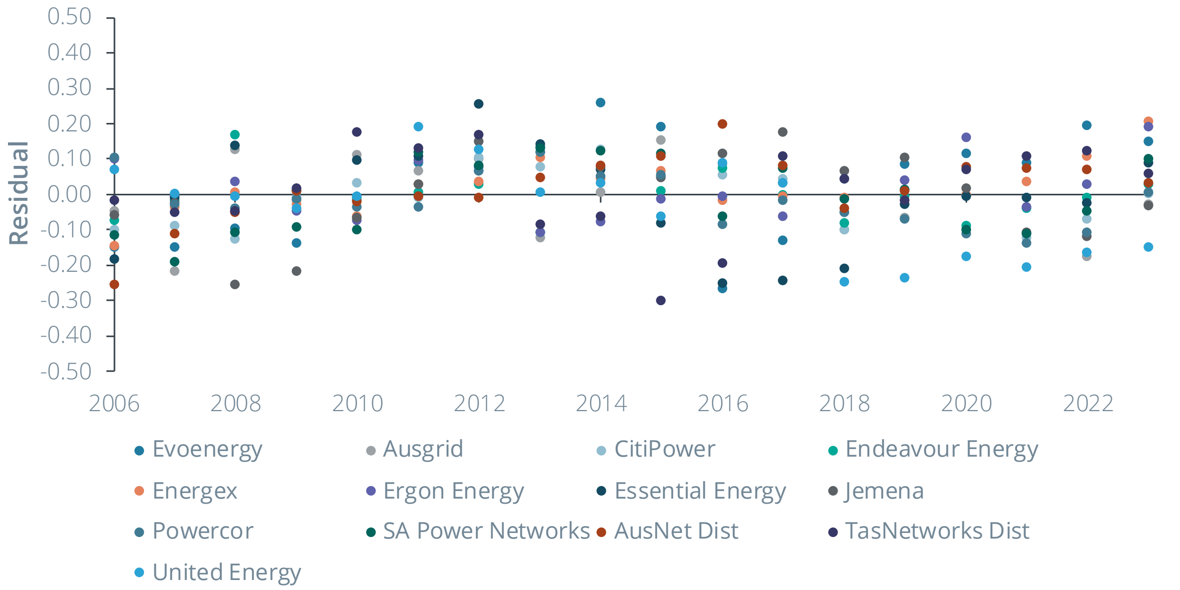
Source: Frontier Economics.

Figure 53: Residuals plot over fitted value for Kumb-AGECJUR-HN-TL



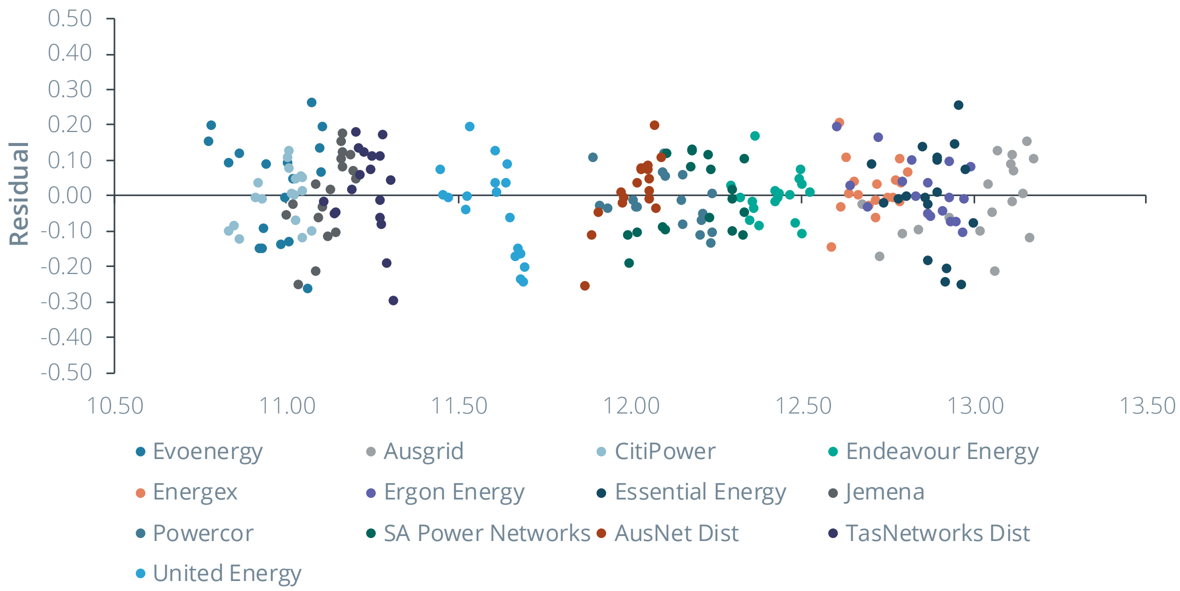
Source: Frontier Economics.

Figure 54: Residuals plot over time for Kumb-JTT-HN-GTC-TL



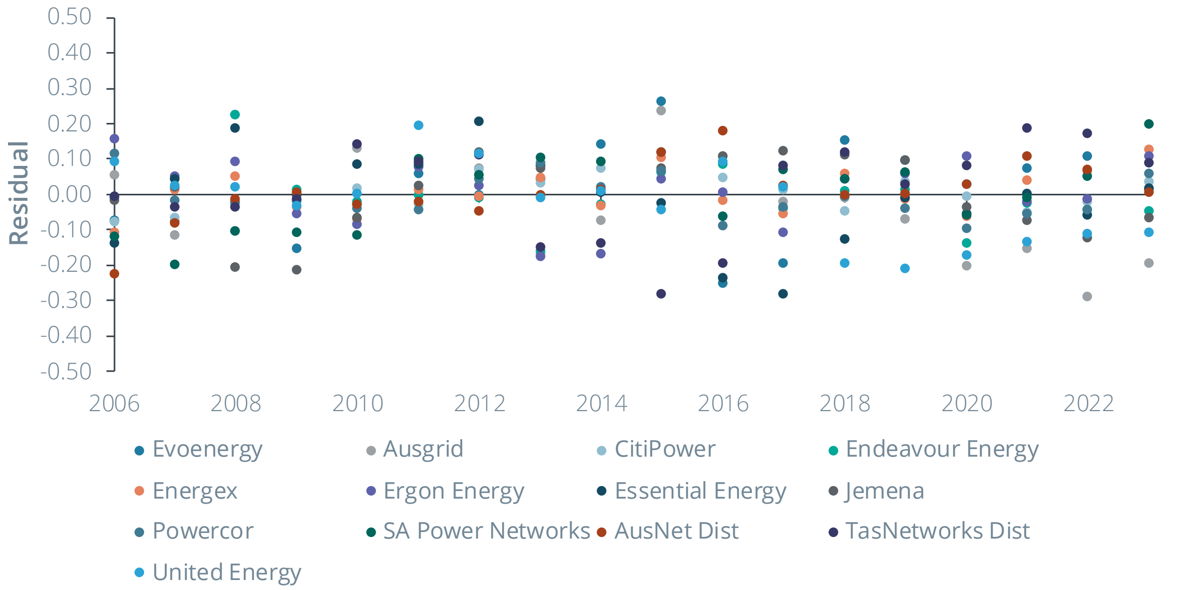
Source: Frontier Economics.

Figure 55: Residuals plot over fitted value for Kumb-JTT-HN-GTC-TL



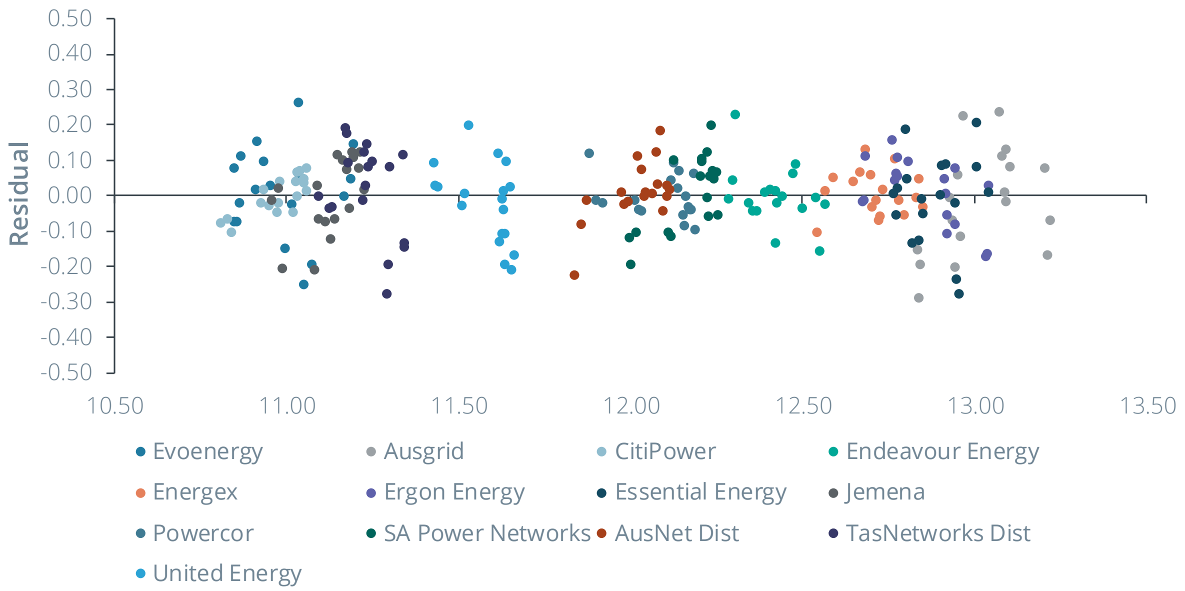
Source: Frontier Economics.

Figure 56: Residuals plot over time for Kumb-AGEC-HN-GTC-TL



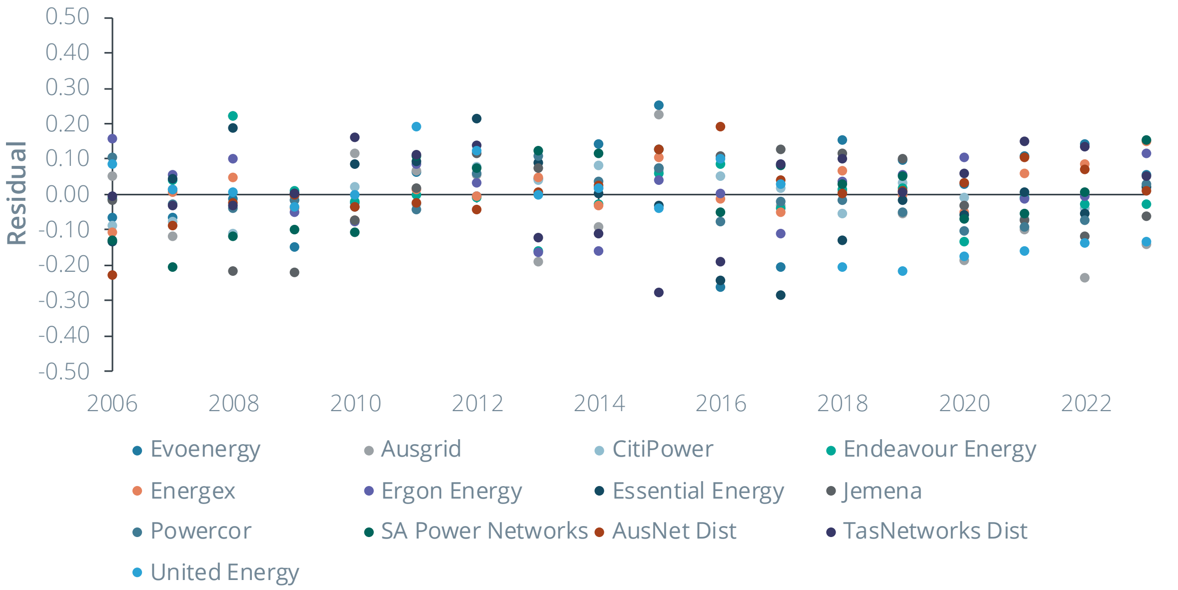
Source: Frontier Economics.

Figure 57: Residuals plot over fitted value for Kumb-AGEC-HN-GTC-TL



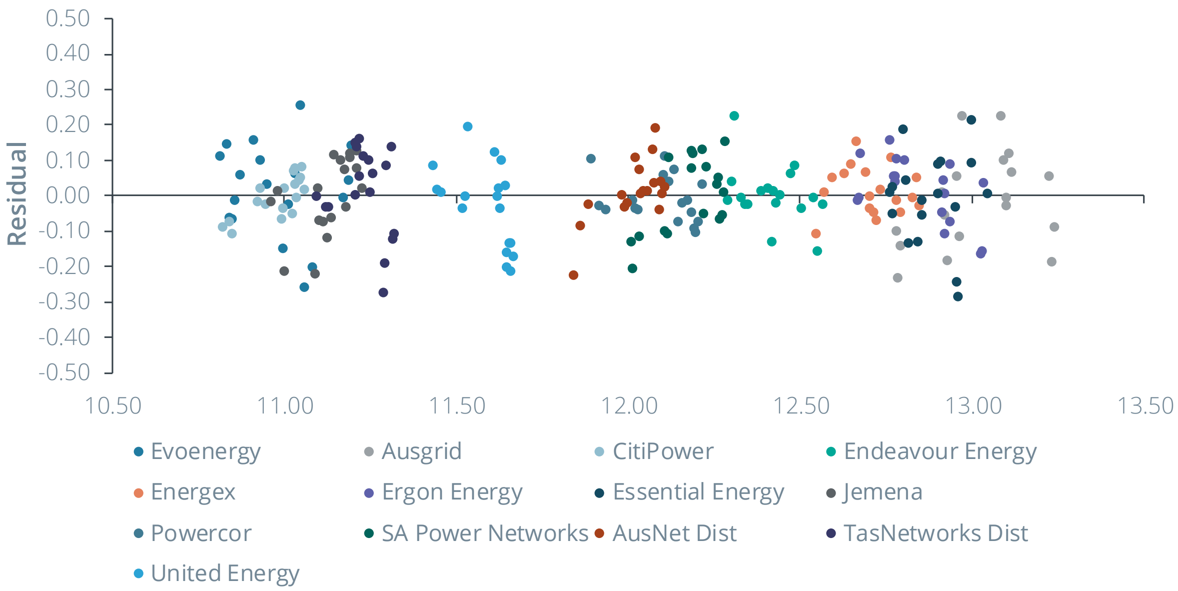
Source: Frontier Economics.

Figure 58: Residuals plot over time for Kumb-AGECJUR-HN-GTC-TL



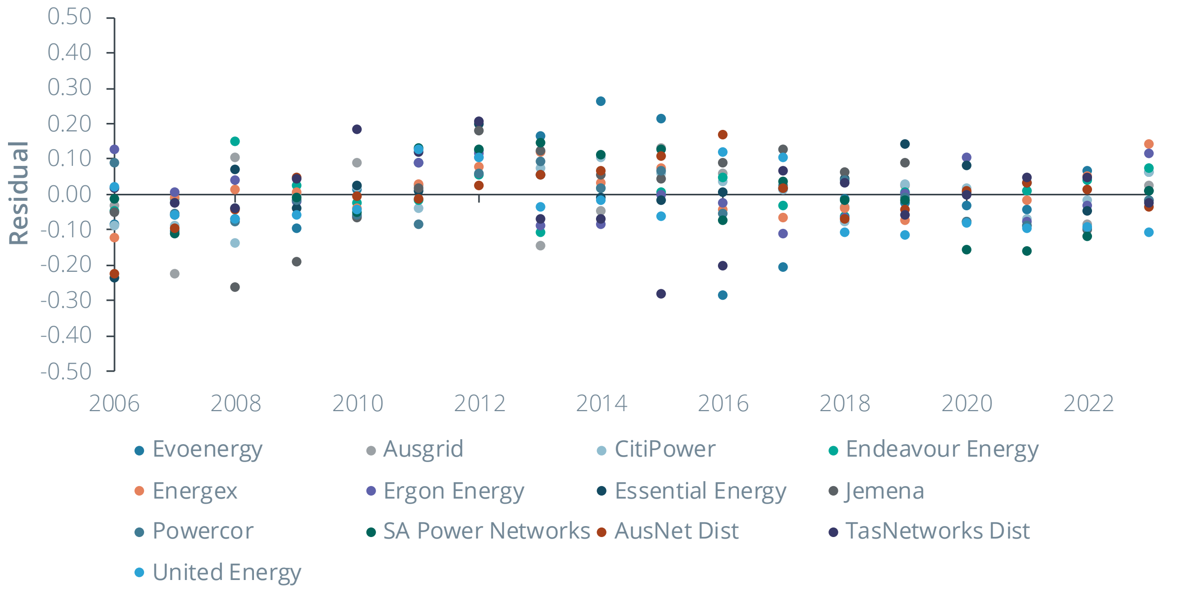
Source: Frontier Economics.

Figure 59: Residuals plot over fitted value for Kumb-AGECJUR-HN-GTC-TL



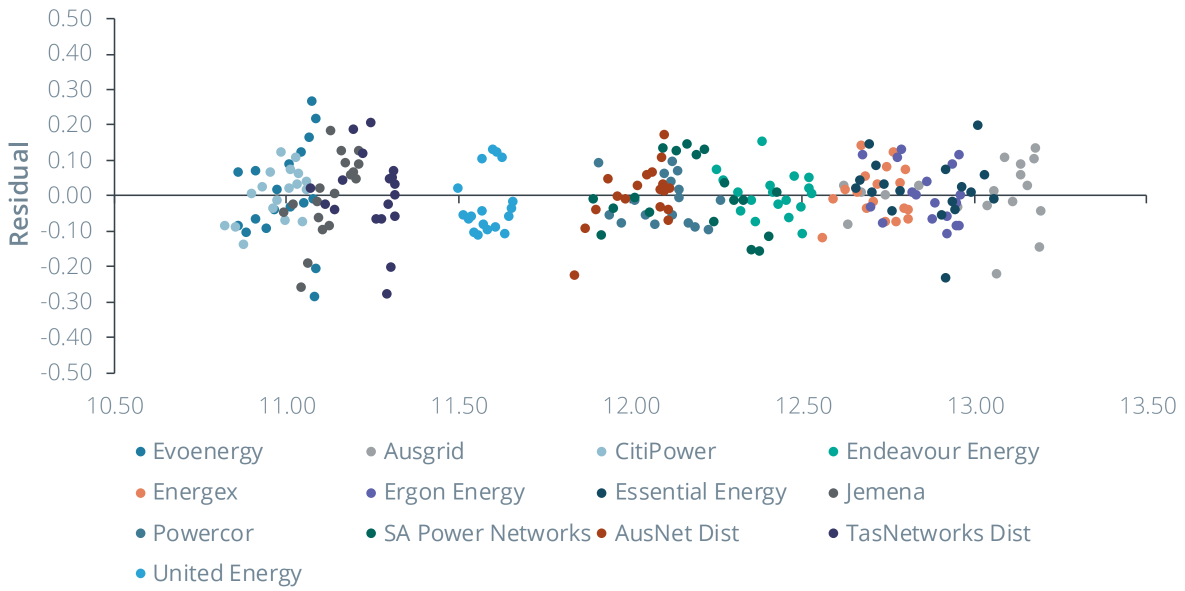
Source: Frontier Economics.

Figure 60: Residuals plot over time for Kumb-AJTT-HN-TL



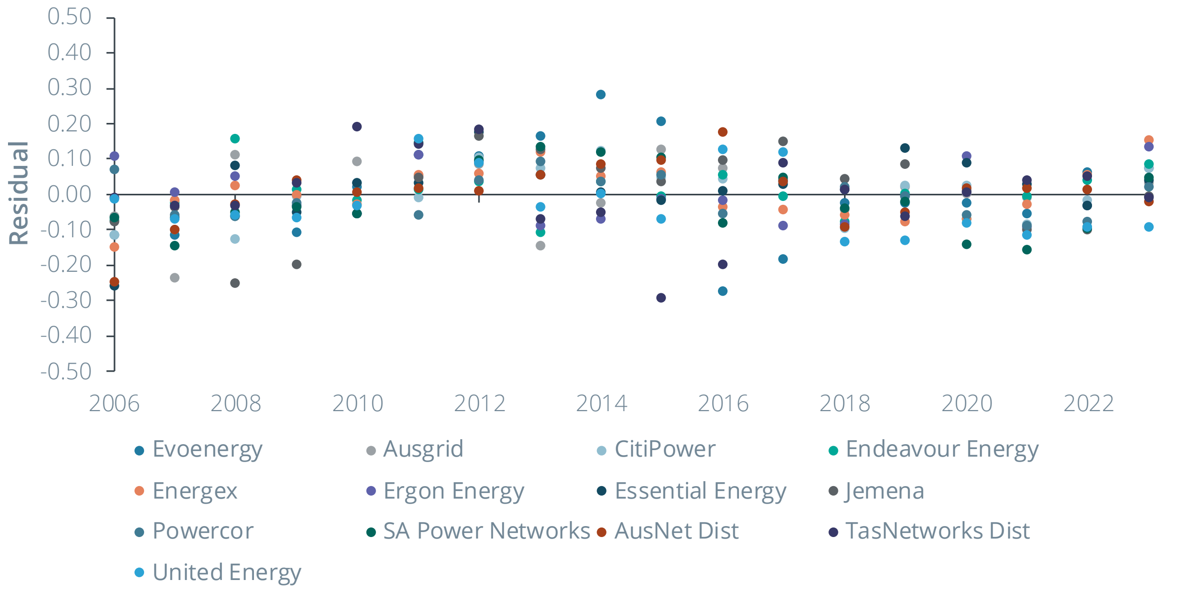
Source: Frontier Economics.

Figure 61: Residuals plot over fitted value for Kumb-AJTT-HN-TL



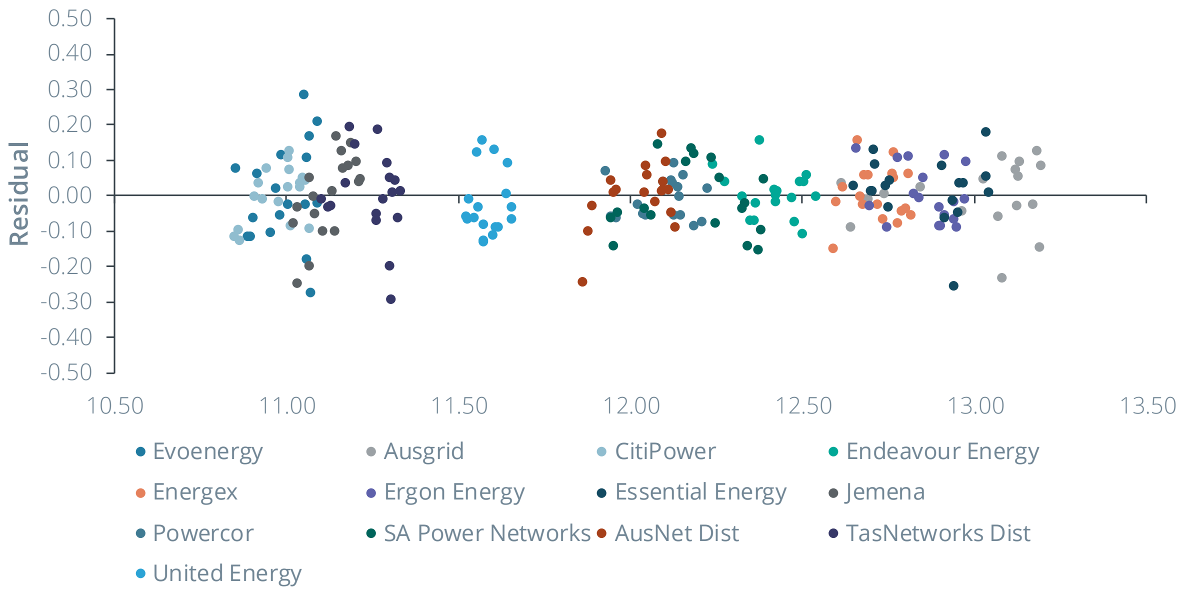
Source: Frontier Economics.

Figure 62: Residuals plot over time for Kumb-AJTT-HN-GTC-TL



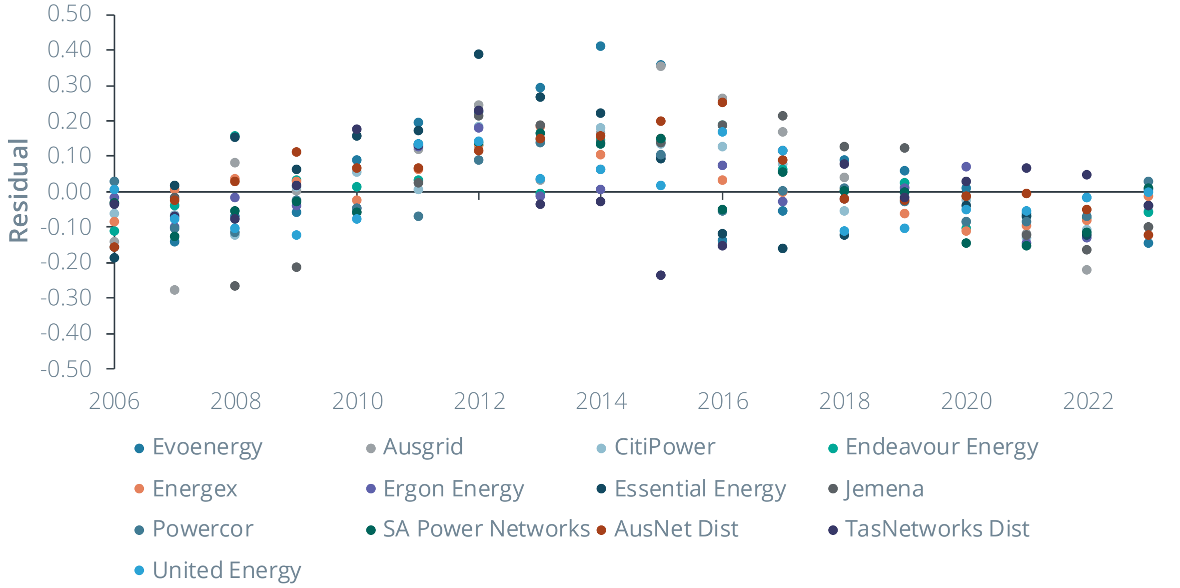
Source: Frontier Economics.

Figure 63: Residuals plot over fitted value for Kumb-AJTT-HN-GTC-TL



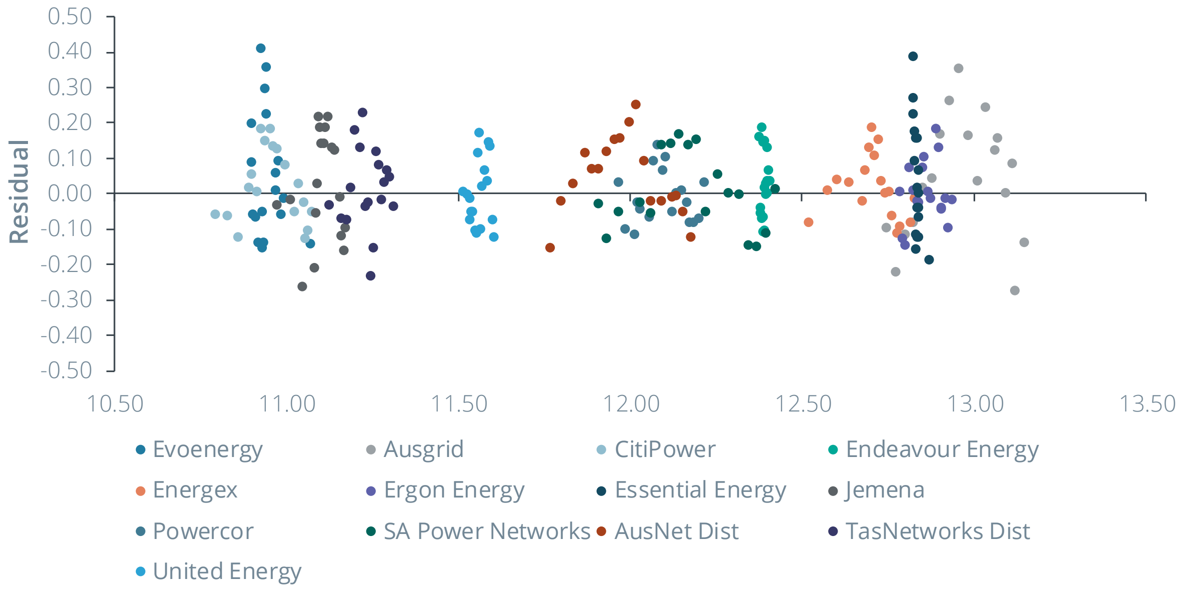
Source: Frontier Economics.

Figure 64: Residuals plot over time for LSE-ADTT-TL



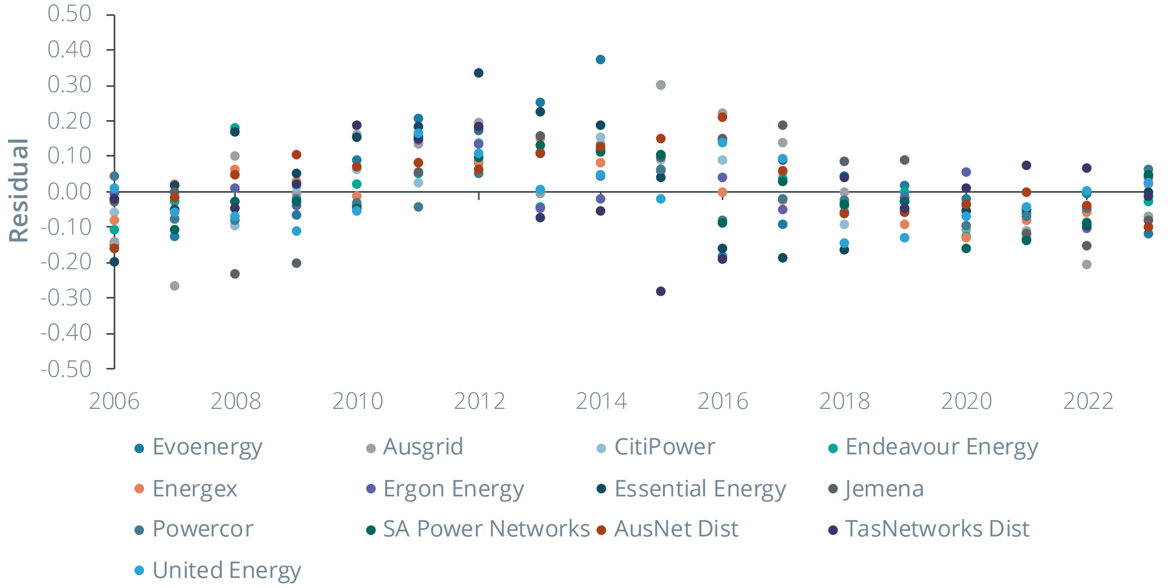
Source: Frontier Economics.

Figure 65: Residuals plot over fitted value for LSE-ADTT-TL



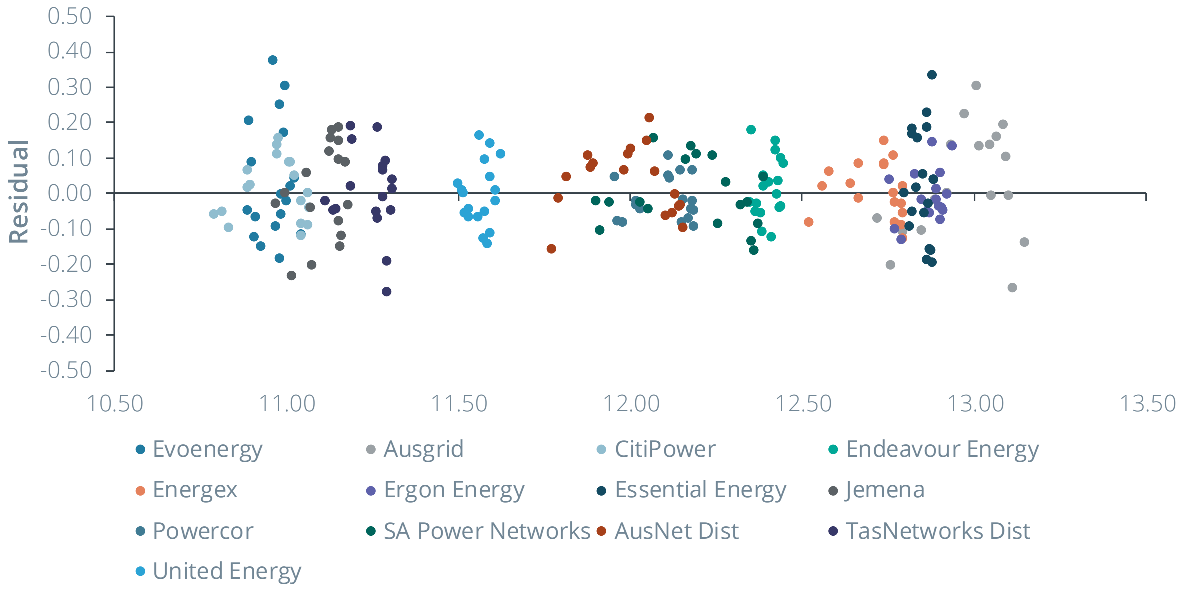
Source: Frontier Economics.

Figure 66: Residuals plot over time for LSE-ADTT-GTC-TL



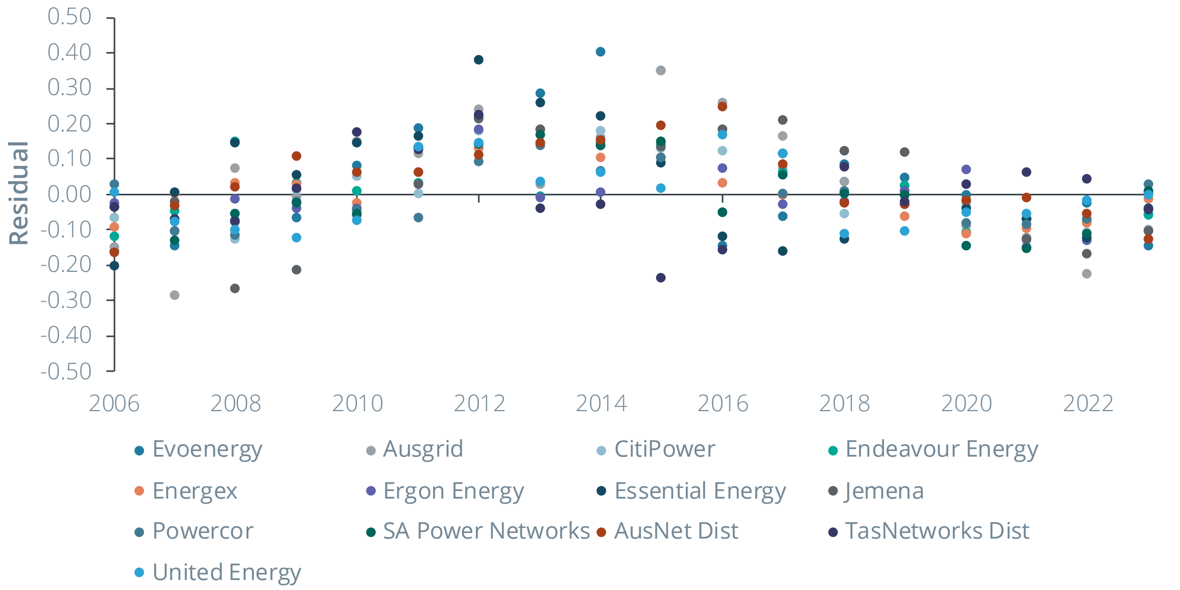
Source: Frontier Economics.

Figure 67: Residuals plot over fitted value for LSE-ADTT-GTC-TL



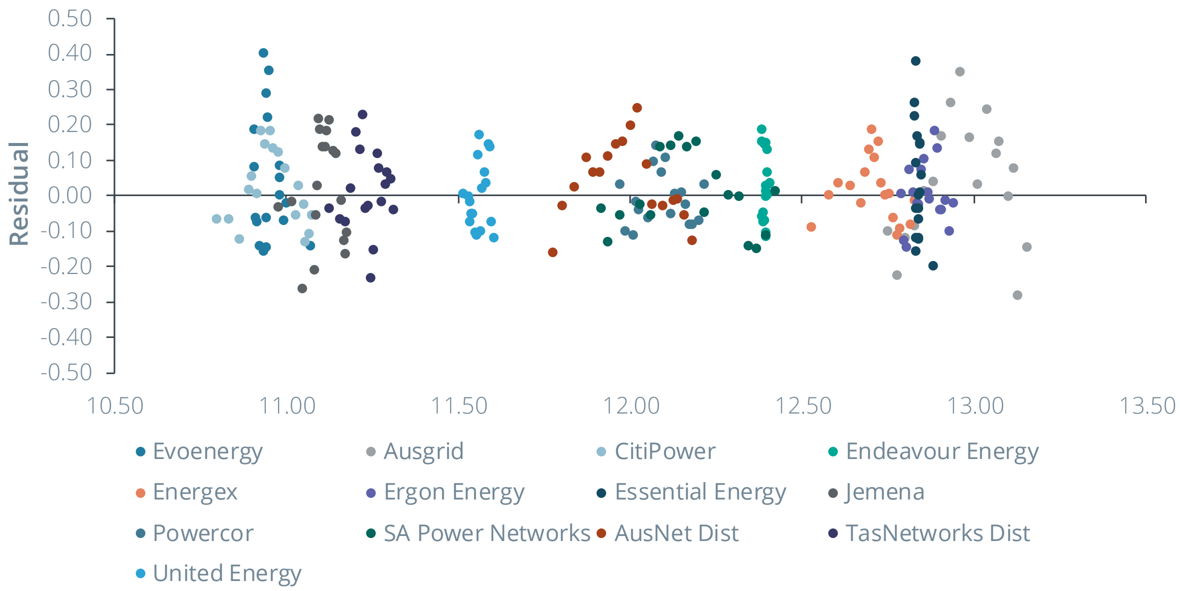
Source: Frontier Economics.

Figure 68: Residuals plot over time for LSE-AJTT-NZ-TL



Source: Frontier Economics.

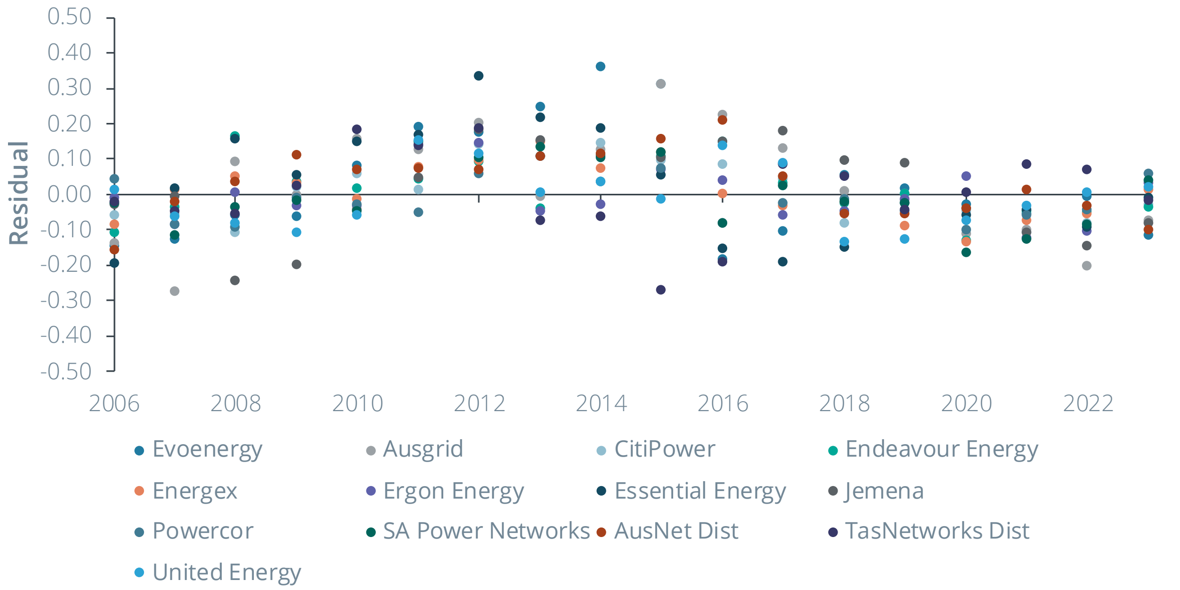
Figure 69: Residuals plot over fitted value for LSE-AJTT-NZ-TL



Source: Frontier Economics.

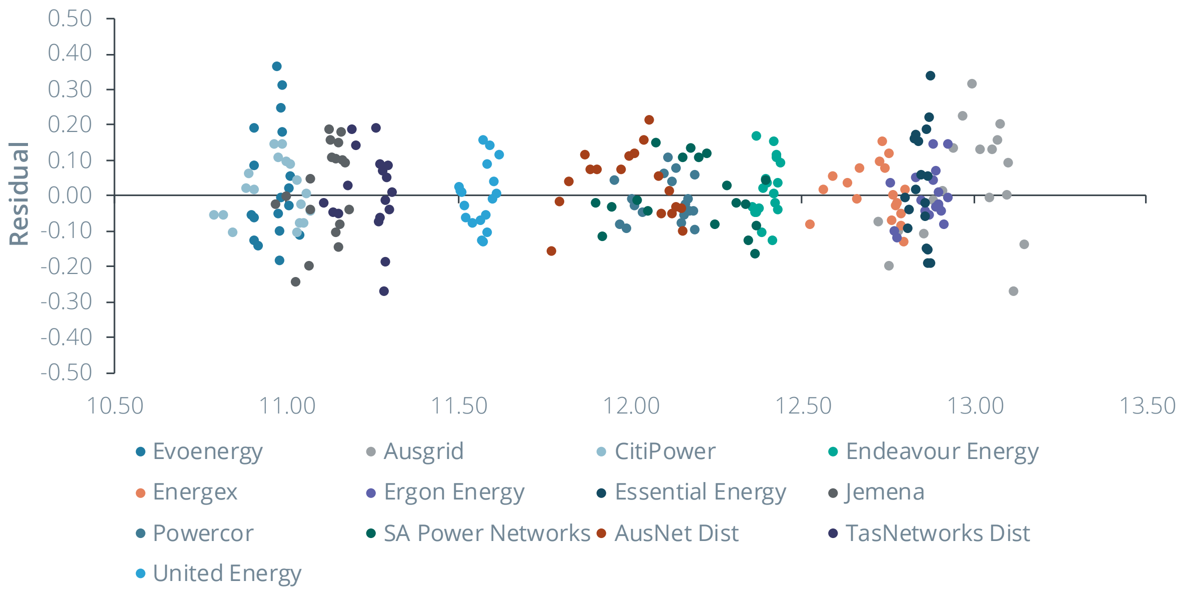
Note: The residuals and coefficients estimate for the New Zealand and Ontario versions of this model are identical, with exception only being the overall year trend. This raises questions of how the roll-forward model should be applied, if the time trend varies depending on the country chosen.

Figure 70: Residuals plot over time for LSE-AJTT-NZ-GTC-TL



Source: Frontier Economics.

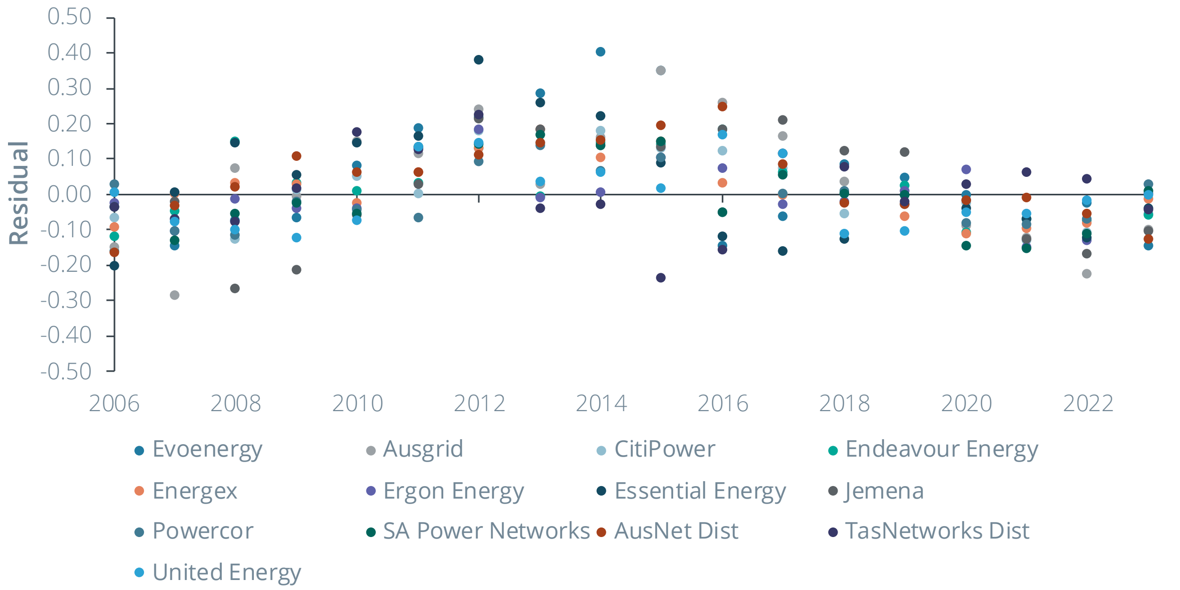
Figure 71: Residuals plot over fitted value for LSE-AJTT-NZ-GTC-TL



Source: Frontier Economics.

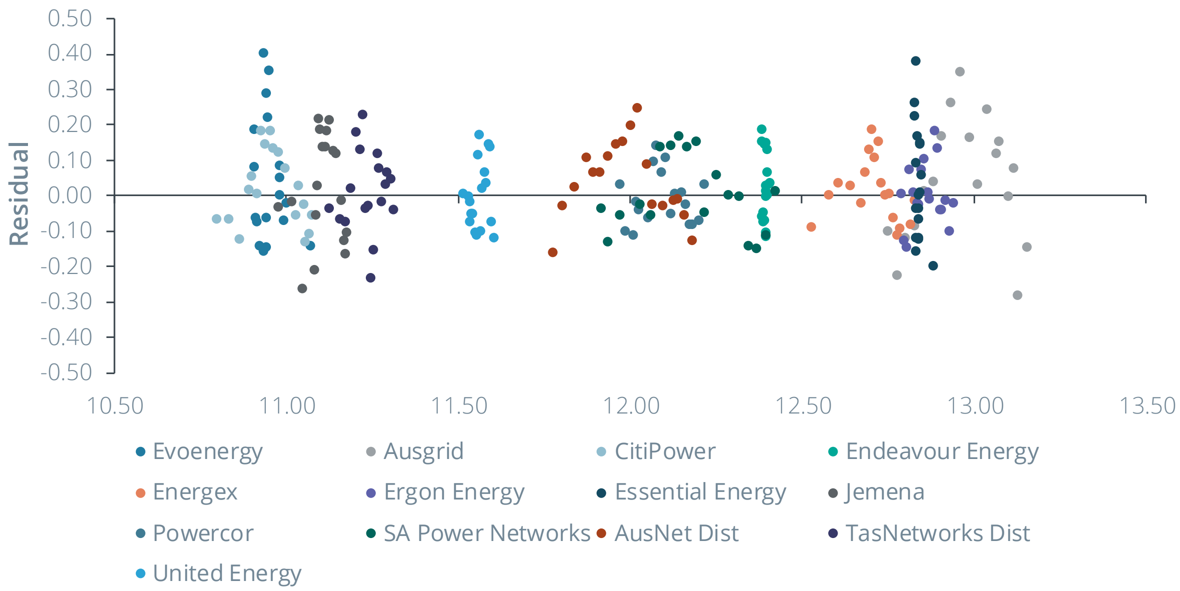
Note: Similarly to the non-GTC versions, The residuals and coefficients estimate for the New Zealand and Ontario versions of this model are very similar. This presents the same interpretation issue for the roll-forward model.

Figure 72: Residuals plot over time for LSE-AJTT-ONT-TL



Source: Frontier Economics.

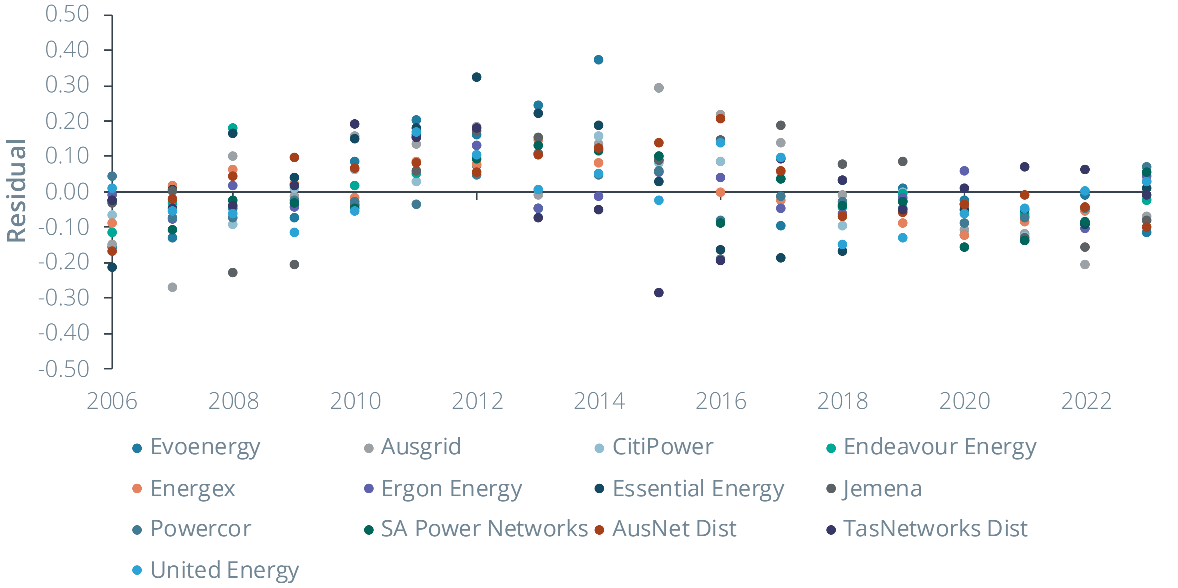
Figure 73: Residuals plot over fitted value for LSE-AJTT-ONT-TL



Source: Frontier Economics.

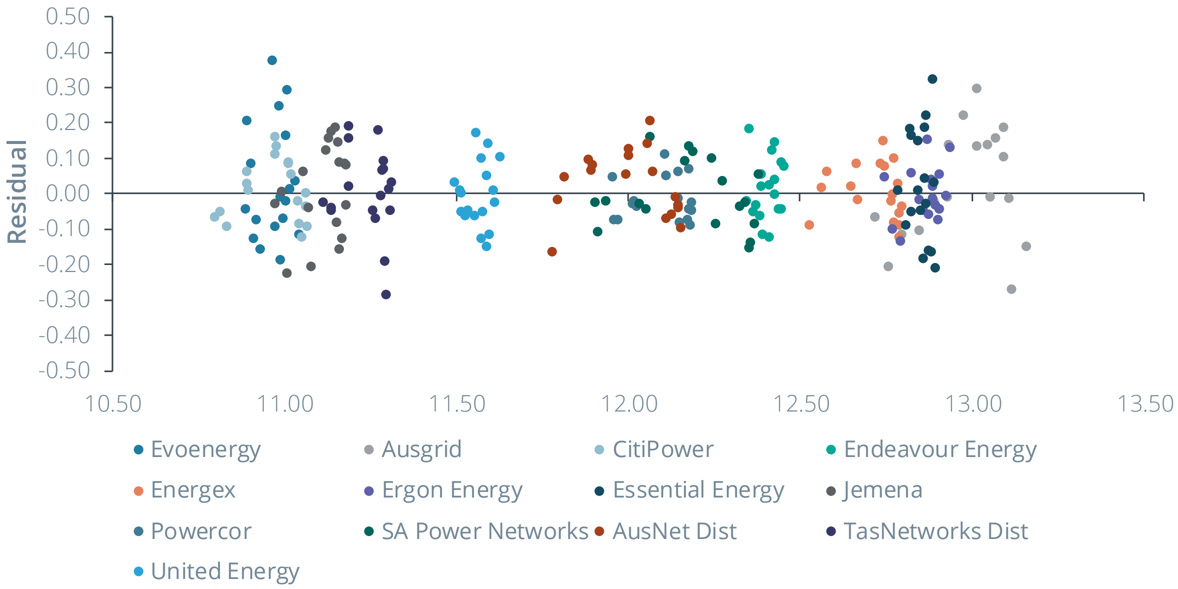
Note: The residuals and coefficients estimate for the New Zealand and Ontario versions of this model are identical, with exception only being the overall year trend. This raises questions of how the roll-forward model should be applied, if the time trend varies depending on the country chosen.

Figure 74: Residuals plot over time for LSE-AJTT-ONT-GTC-TL



Source: Frontier Economics.

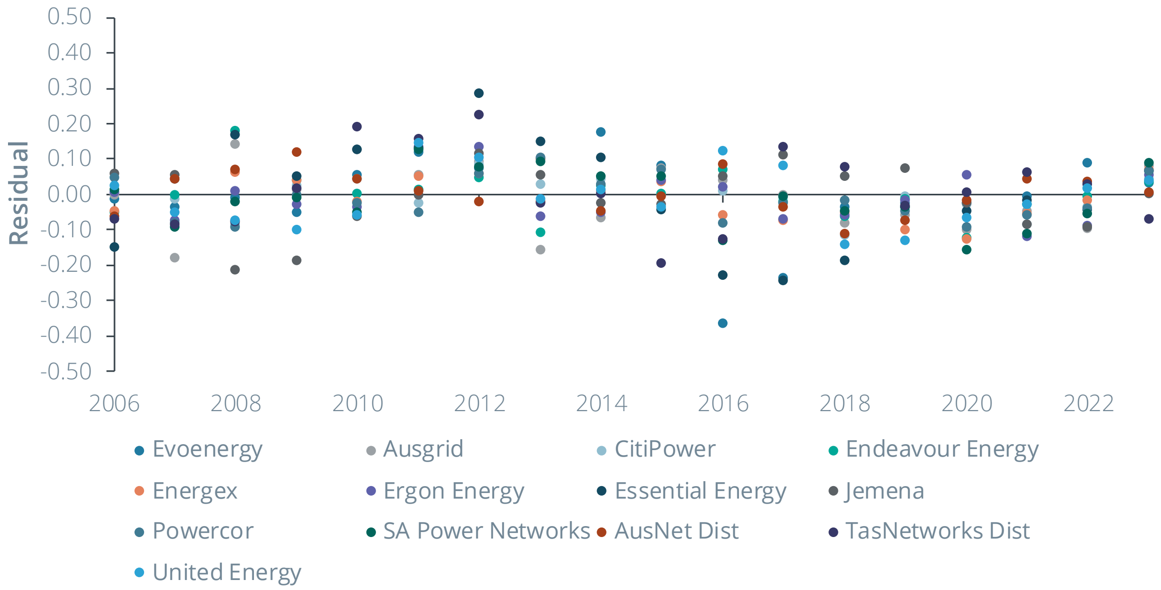
Figure 75: Residuals plot over fitted value for LSE-AJTT-ONT-GTC-TL



Source: Frontier Economics.

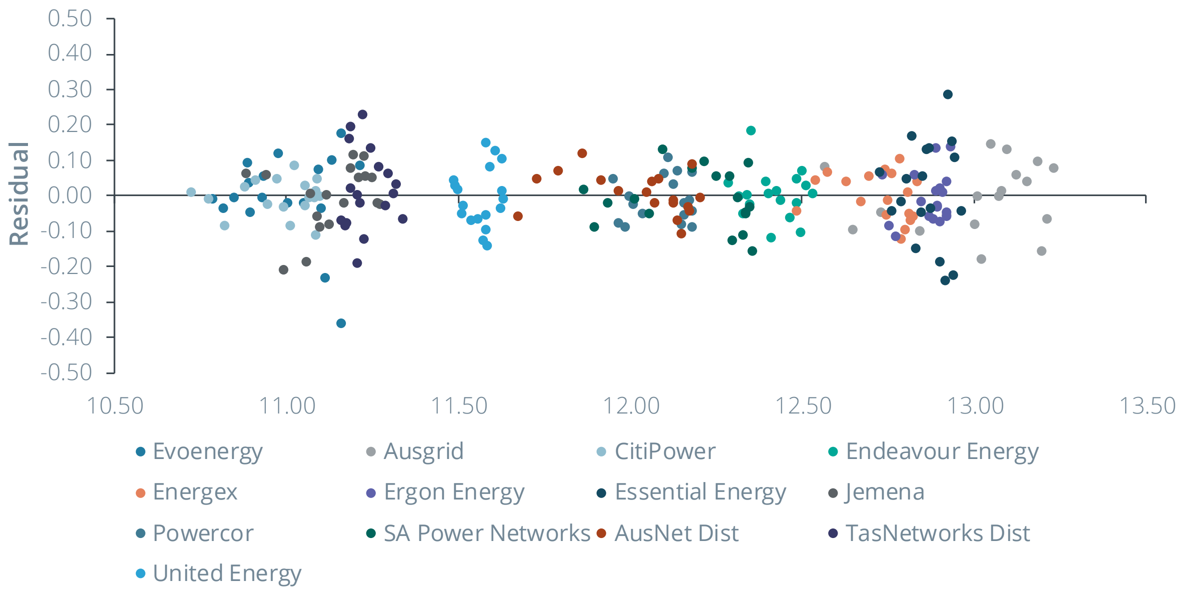
Note: Similarly to the non-GTC versions, The residuals and coefficients estimate for the New Zealand and Ontario versions of this model are very similar. This presents the same interpretation issue for the roll-forward model.

Figure 76: Residuals plot over time for LSE-AJTTBREAK-NZ-GTC-TL



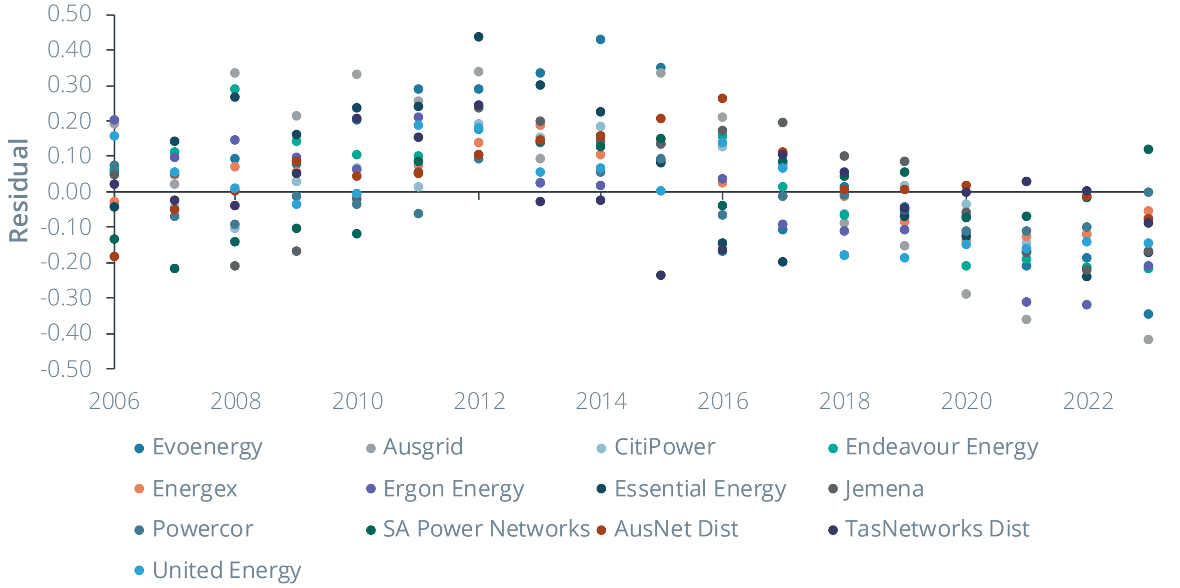
Source: Frontier Economics.

Figure 77: Residuals plot over fitted value for LSE-AJTTBREAK-NZ-GTC-TL



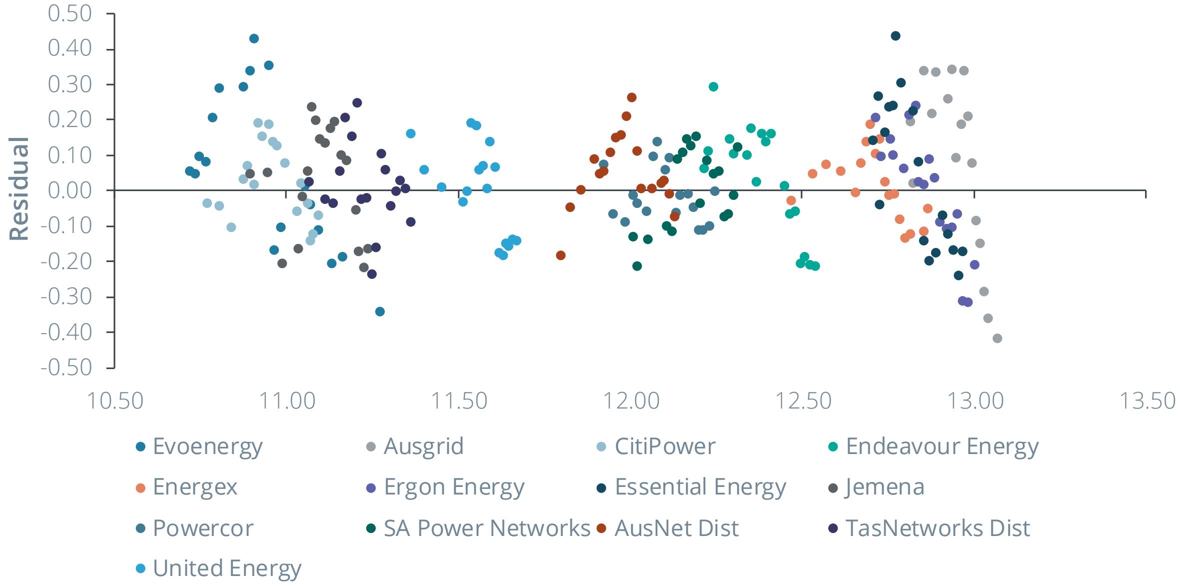
Source: Frontier Economics.

Figure 78: Residuals plot over time for LSE-TI-TL



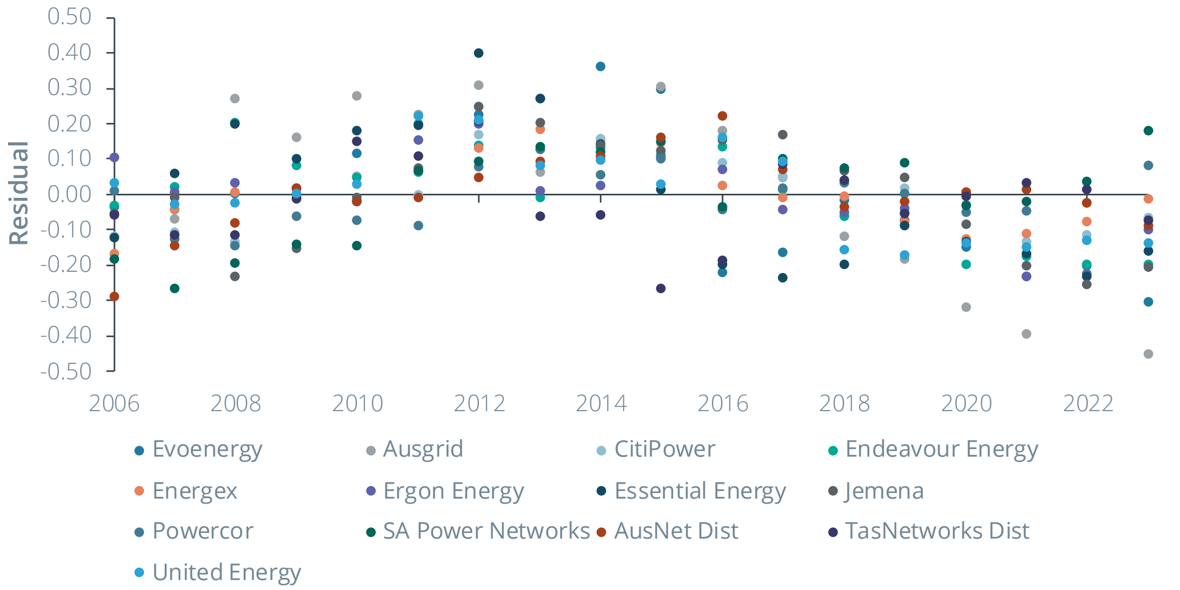
Source: Frontier Economics.

Figure 79: Residuals plot over fitted value for LSE-TI-TL



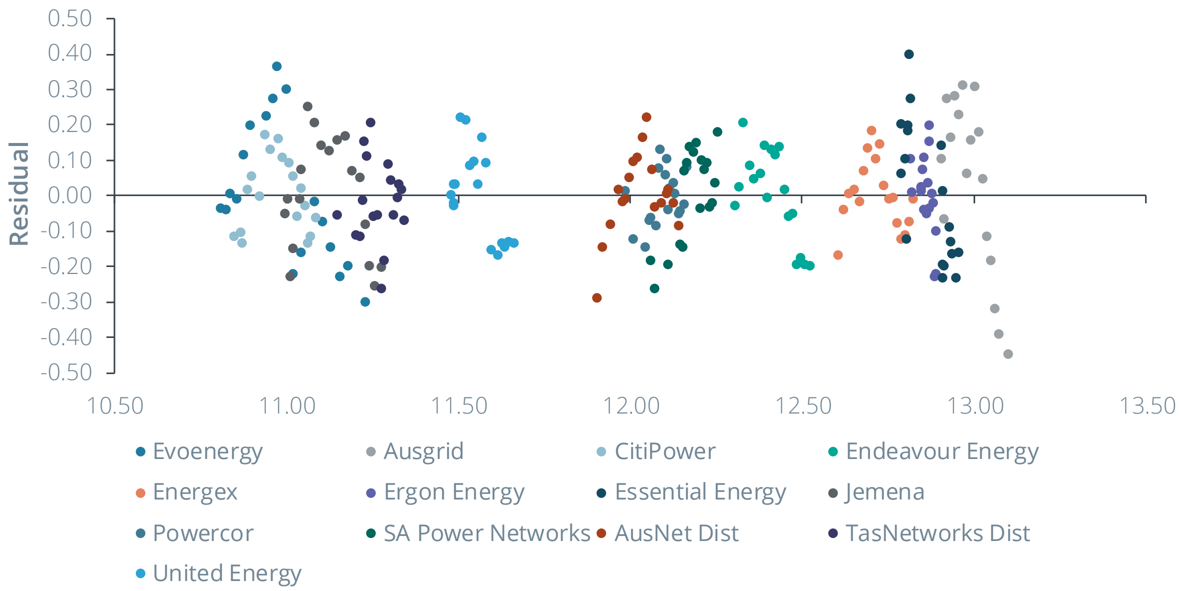
Source: Frontier Economics.

Figure 80: Residuals plot over time for SFA-TI-TL



Source: Frontier Economics.

Figure 81: Residuals plot over fitted value for SFA-TI-TL

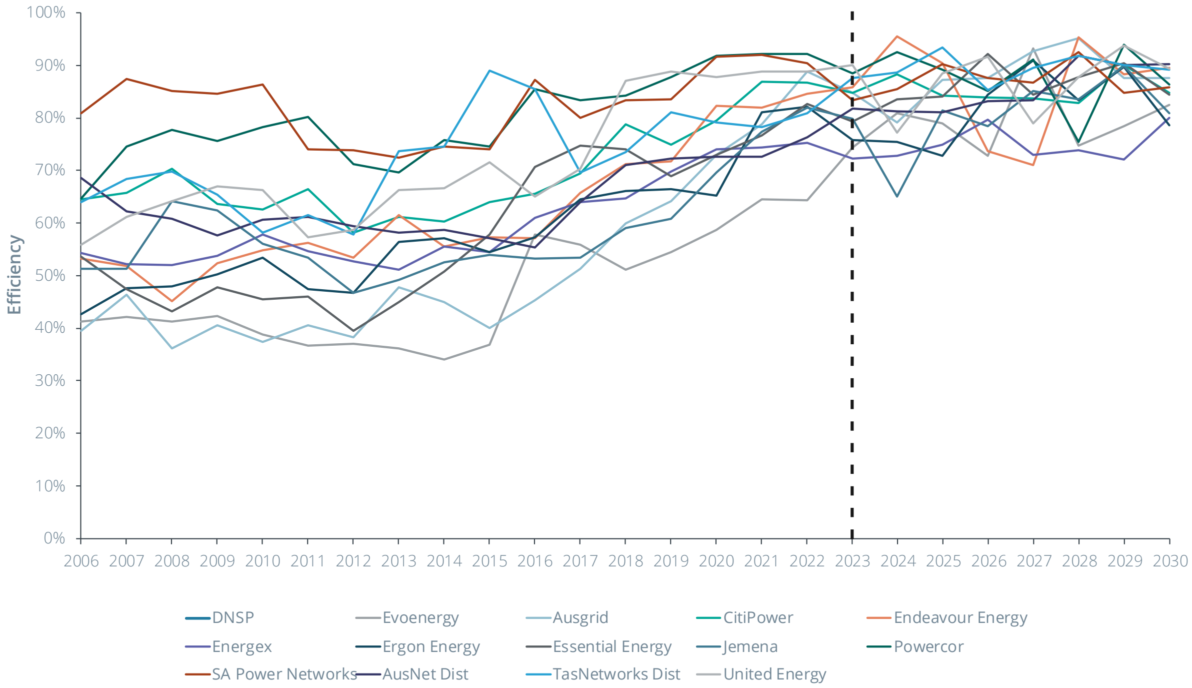


Source: Frontier Economics.

Stability testing

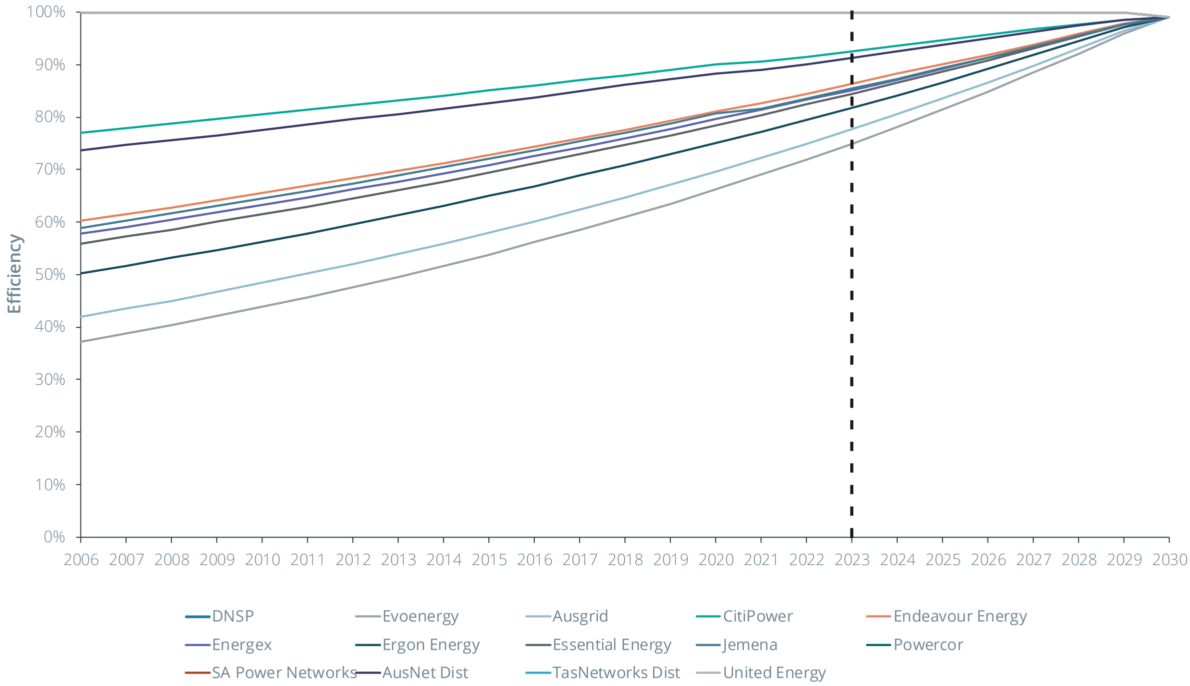
In this section we provide figures detailing the results of our stability testing using generated data with known efficiency scores. Data was generated using the BC95 model, holding DNSP inefficiency constant at the level of 2023 through to 2030. The period of generated data is visually denoted by the black segmented line.

Figure 82: Efficiency predictions under constant efficiency data generation – BC95-JTT-HN-TL



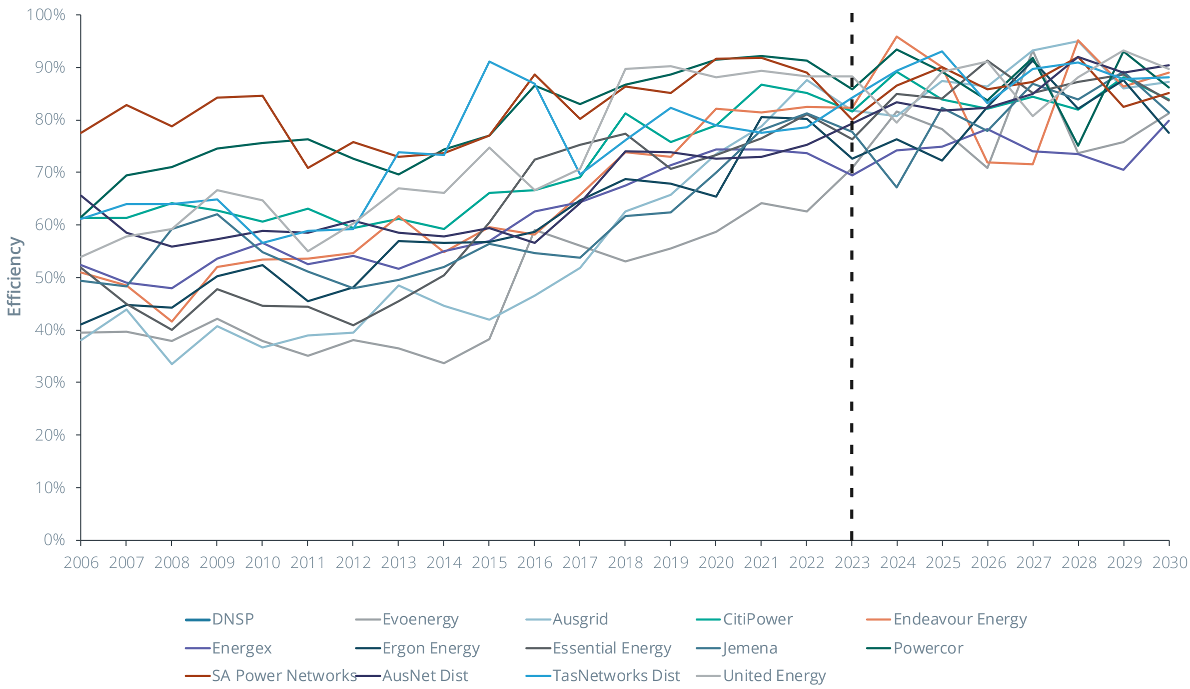
Source: Frontier Economics.

Figure 83: Efficiency predictions under constant efficiency data generation – BC95-AJTT-HN-TL



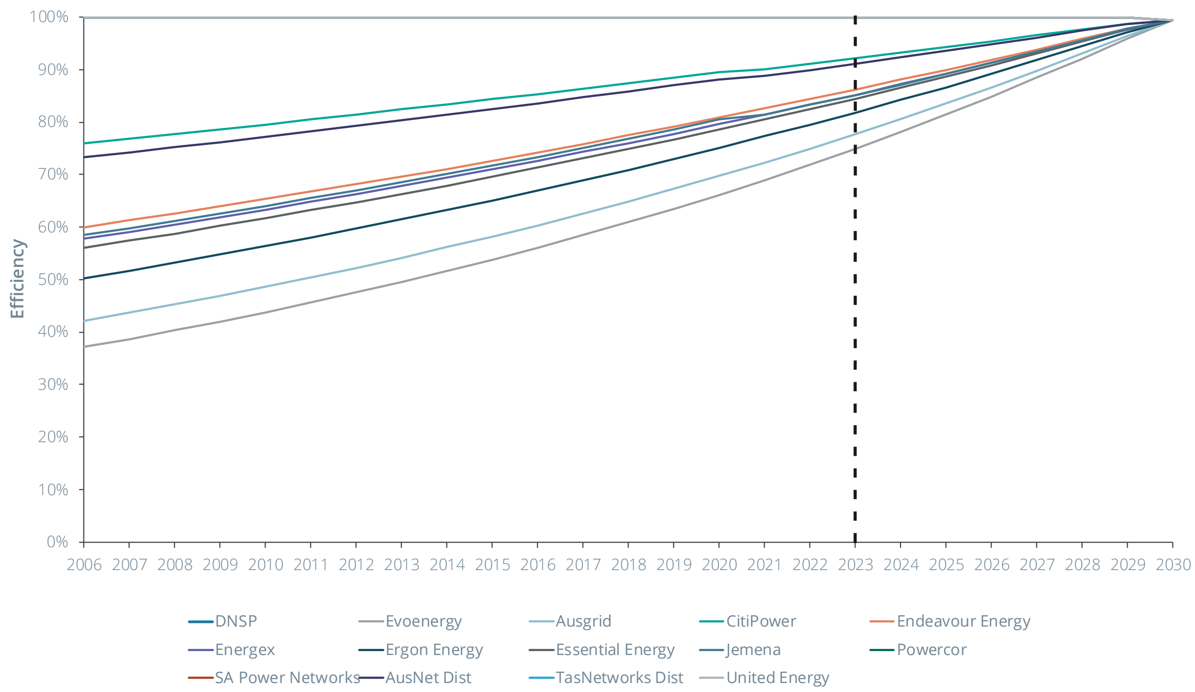
Source: Frontier Economics.

Figure 84: Efficiency predictions under constant efficiency data generation – BC95-JTT-HN-GTC-TL



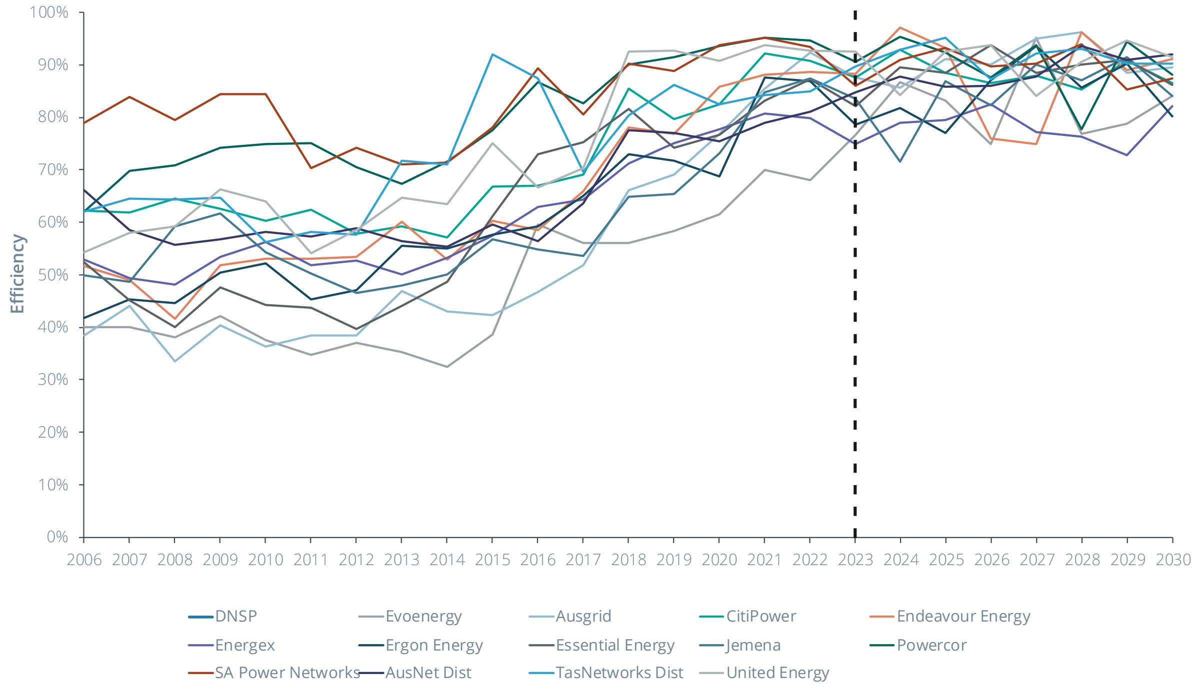
Source: Frontier Economics.

Figure 85: Efficiency predictions under constant efficiency data generation – BC95-AJTT-HN-GTC-TL



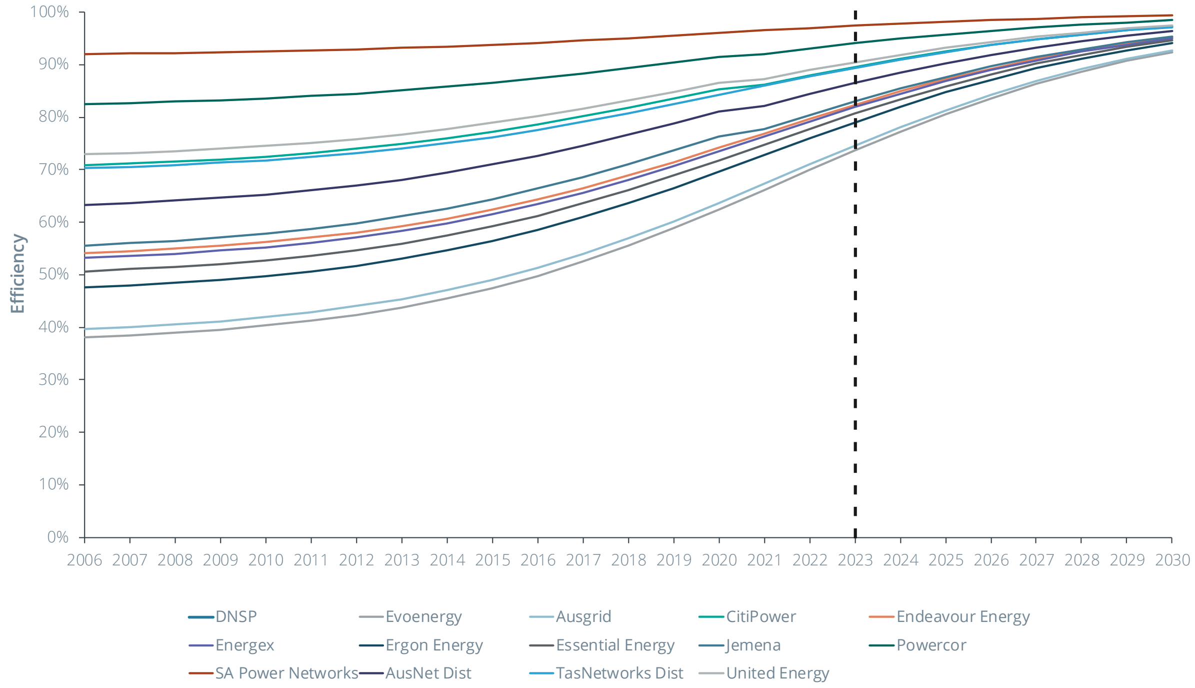
Source: Frontier Economics.

Figure 86: Efficiency predictions under constant efficiency data generation – BC95-AGTT-HN-GTC-TL



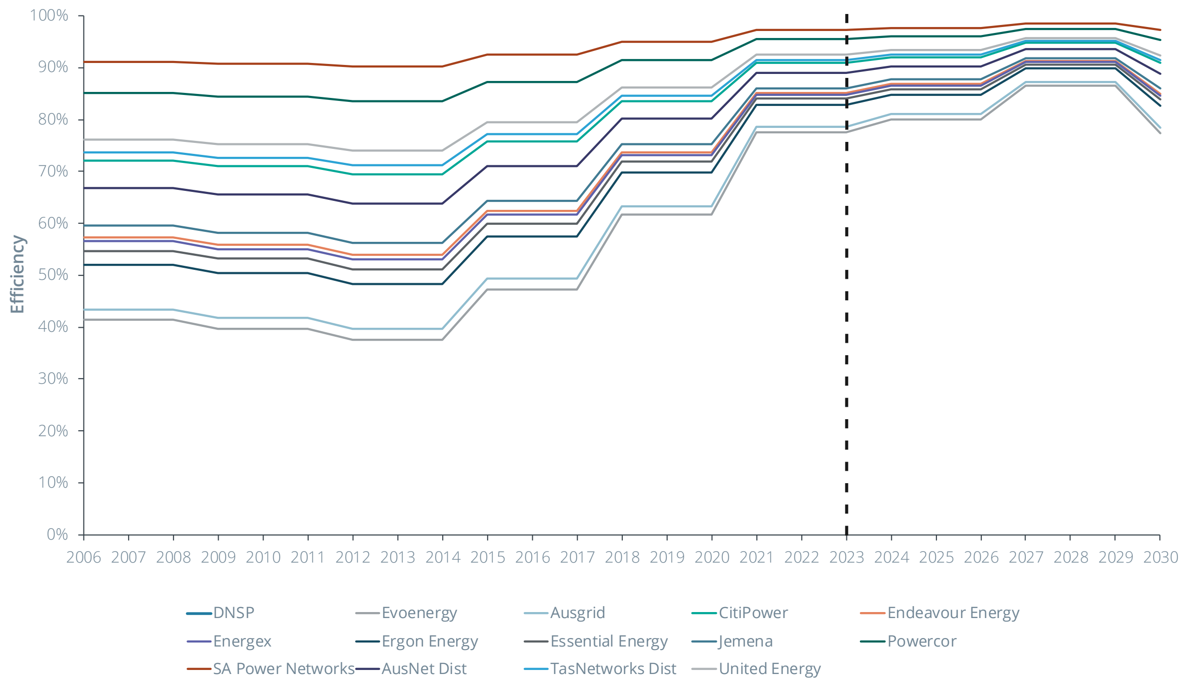
Source: Frontier Economics.

Figure 87: Efficiency predictions under constant efficiency data generation – Kumb-JTT-HN-TL



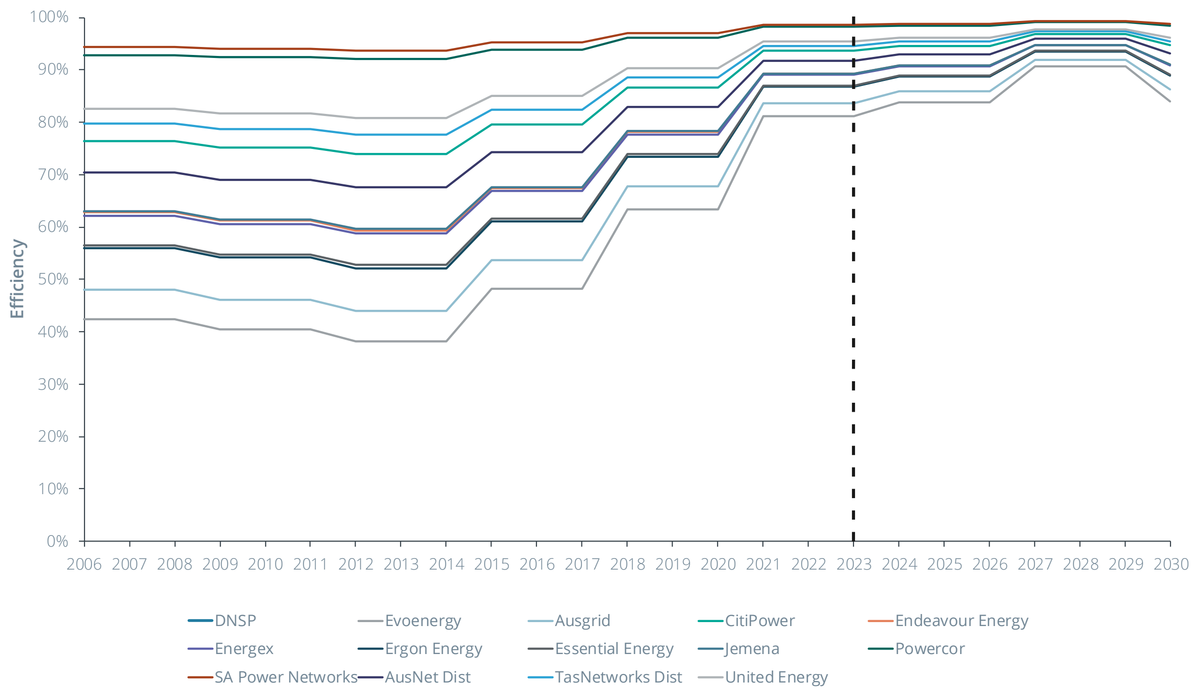
Source: Frontier Economics.

Figure 88: Efficiency predictions under constant efficiency data generation – Kumb-AGTC-HN-TL



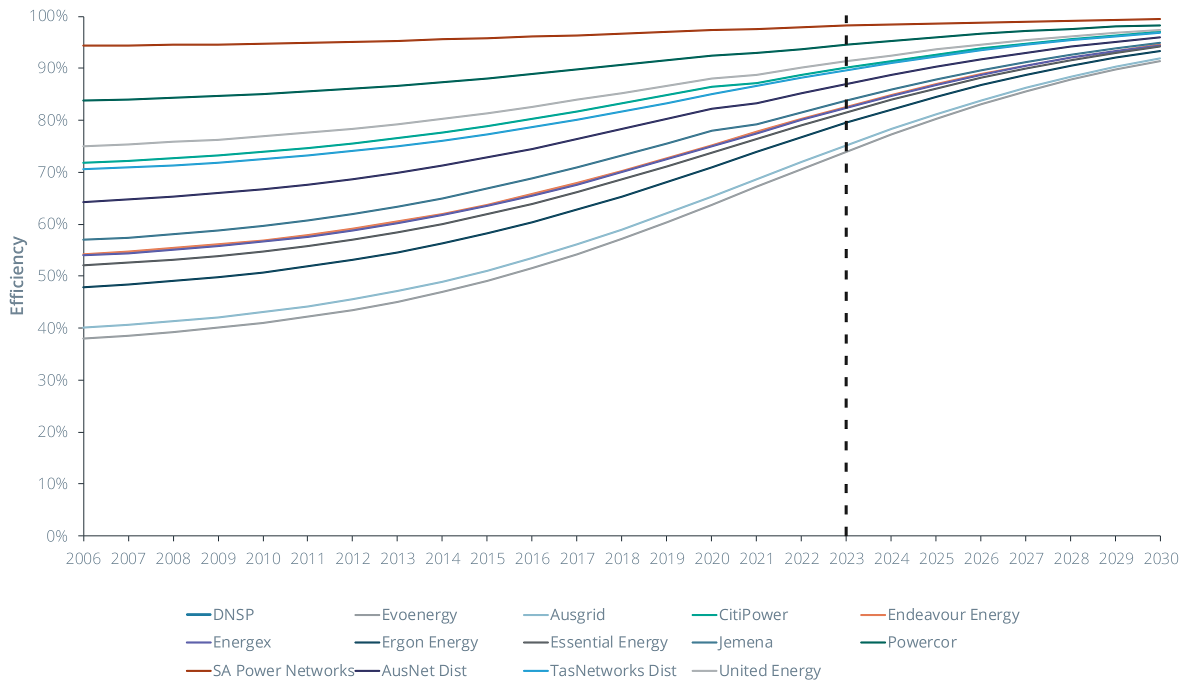
Source: Frontier Economics.

Figure 89: Efficiency predictions under constant efficiency data generation – Kumb-AGTCcon-HN-TL



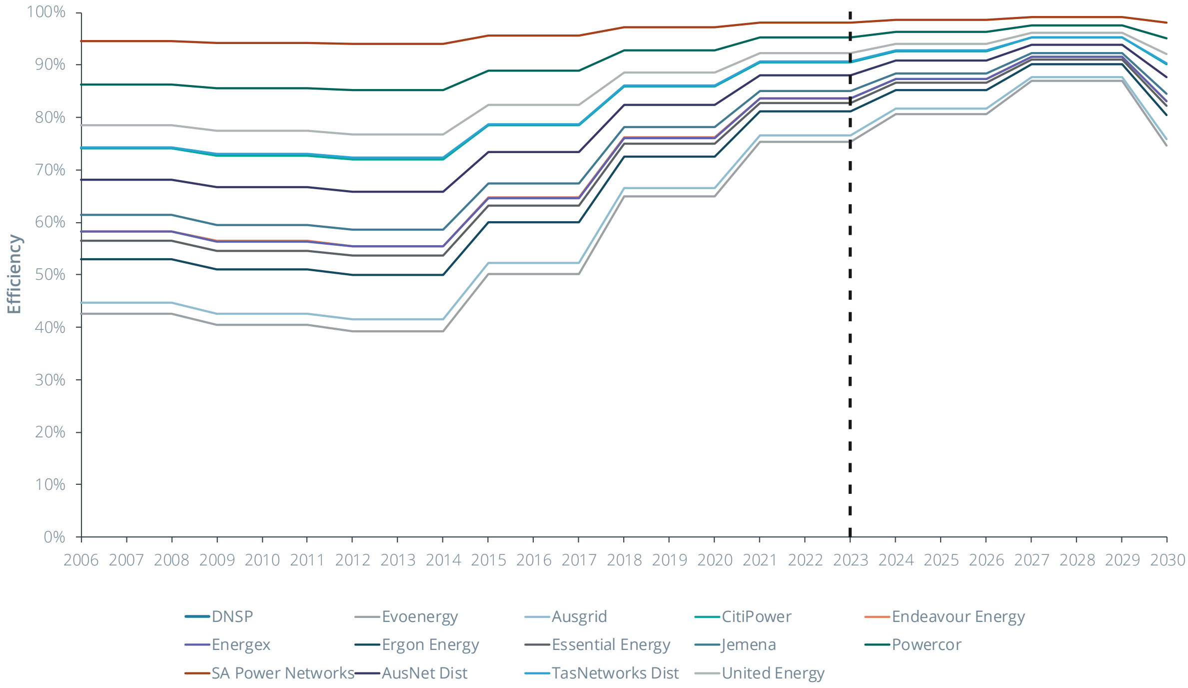
Source: Frontier Economics.

Figure 90: Efficiency predictions under constant efficiency data generation – Kumb-JTT-HN-GTC-TL



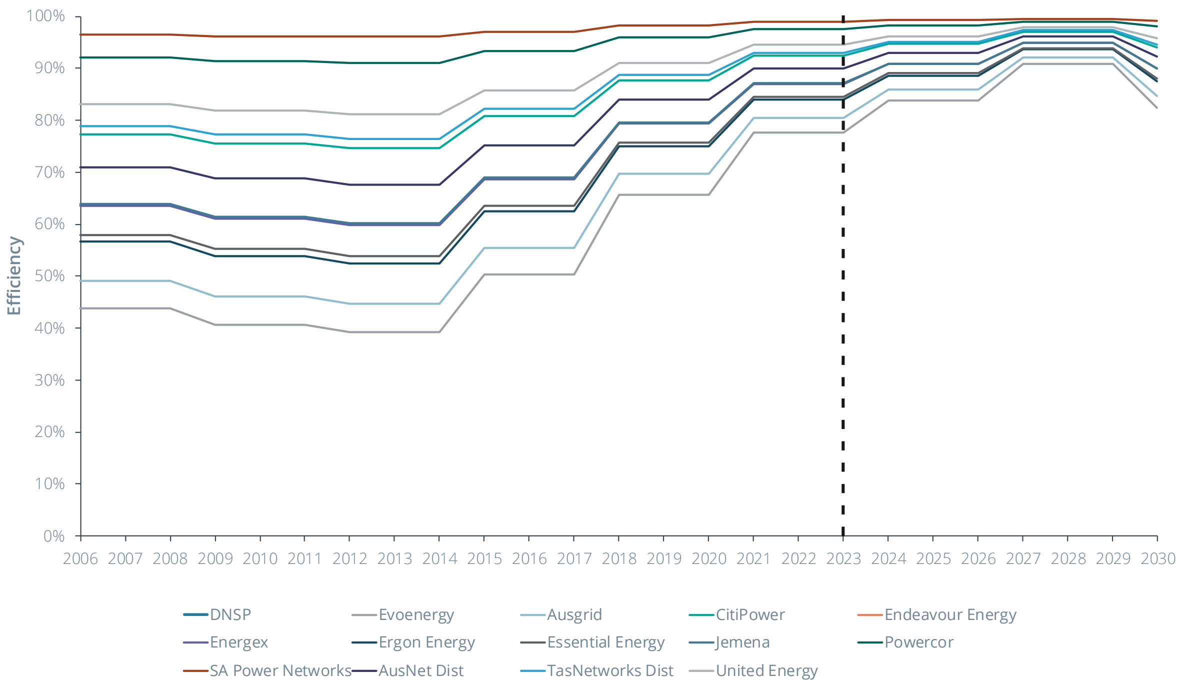
Source: Frontier Economics.

Figure 91: Efficiency predictions under constant efficiency data generation – Kumb-AGTC-HN-GTC-TL



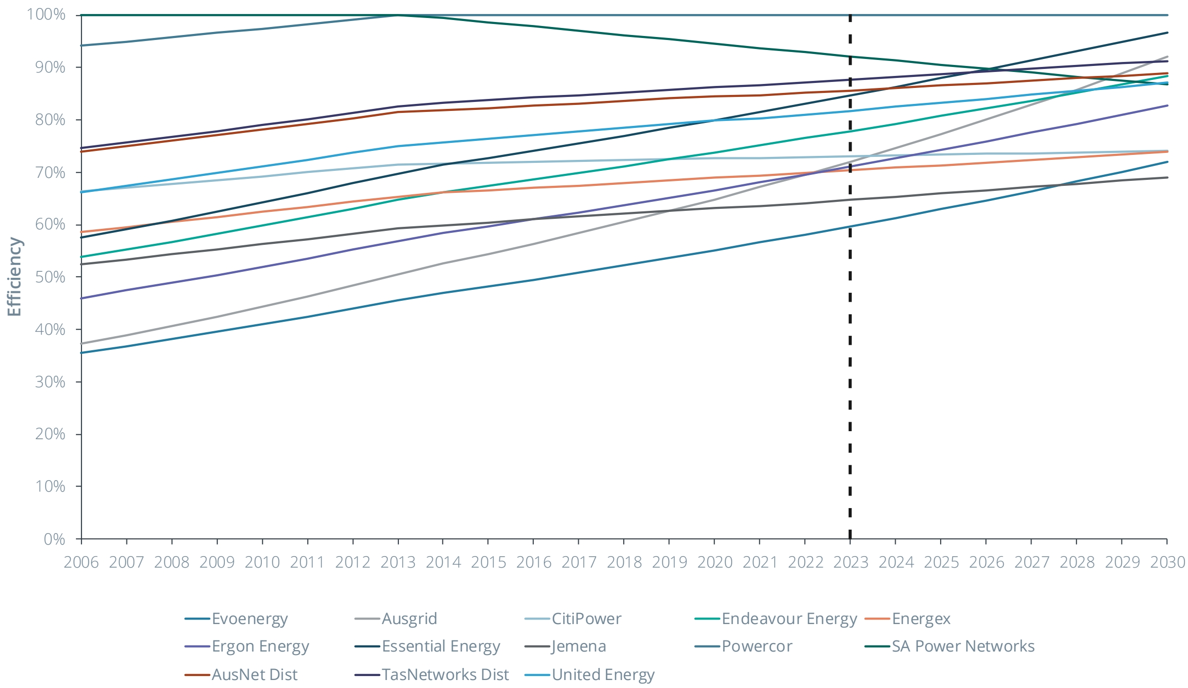
Source: Frontier Economics.

Figure 92: Efficiency predictions under constant efficiency data generation – Kumb-AGTCon-HN-GTC-TL



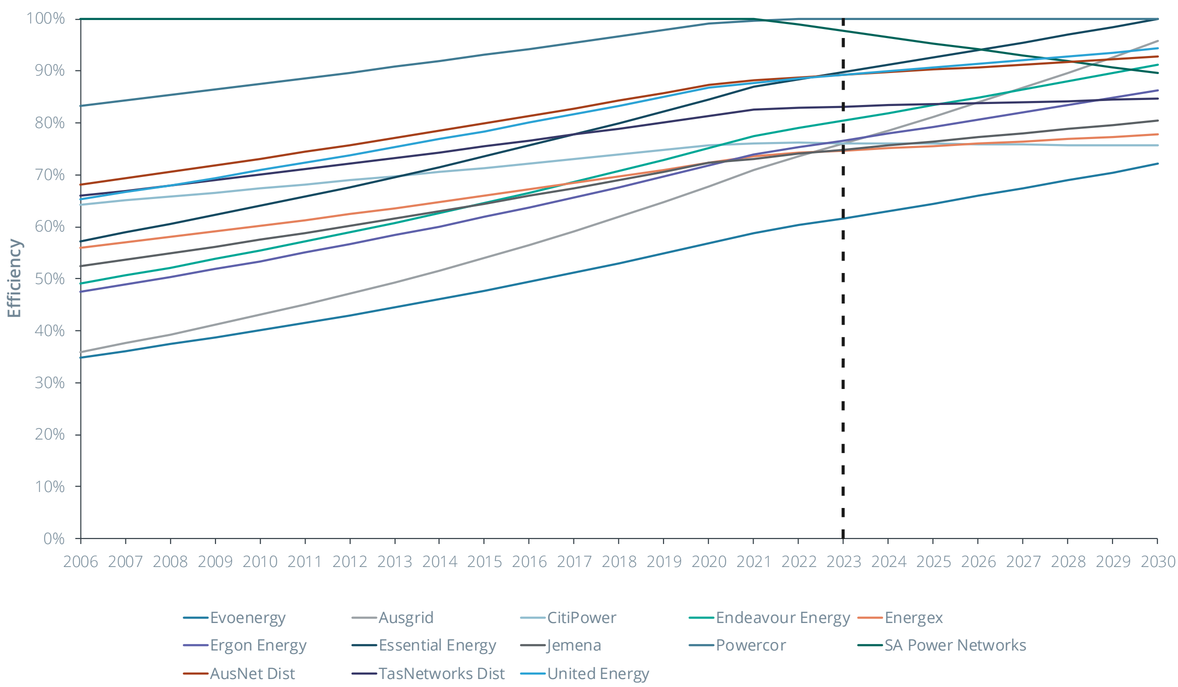
Source: Frontier Economics.

Figure 93: Efficiency predictions under constant efficiency data generation – LSE-ADTT-TL



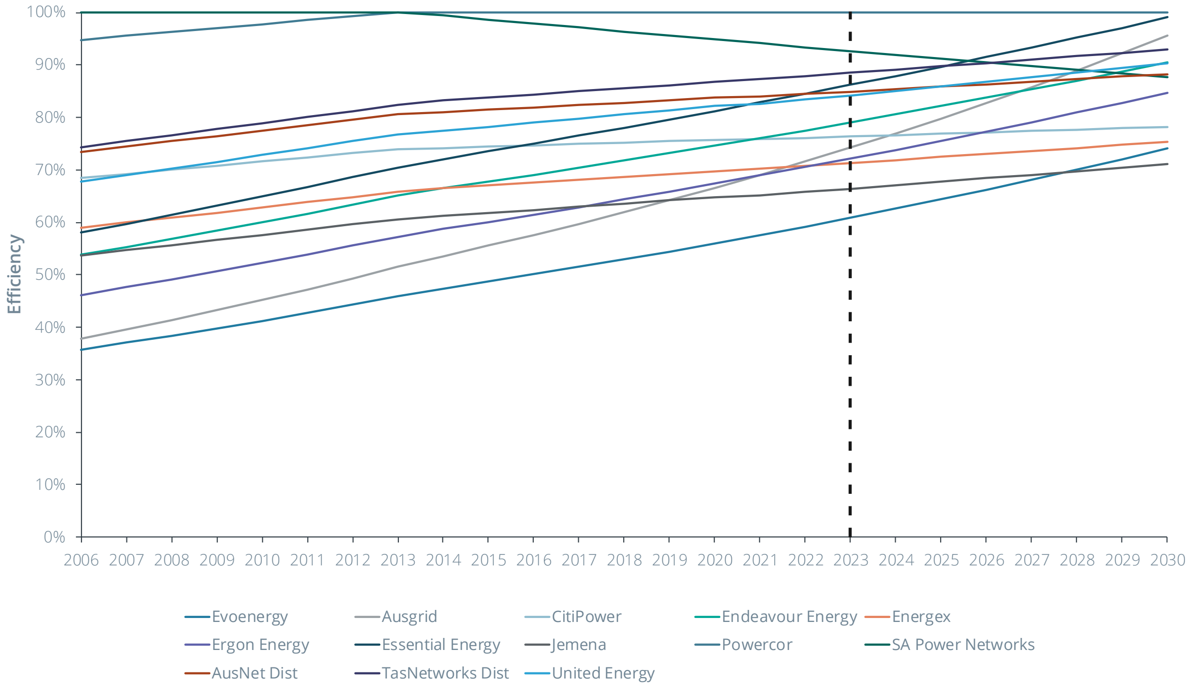
Source: Frontier Economics.

Figure 94: Efficiency predictions under constant efficiency data generation – LSE-ADTT-GTC-TL



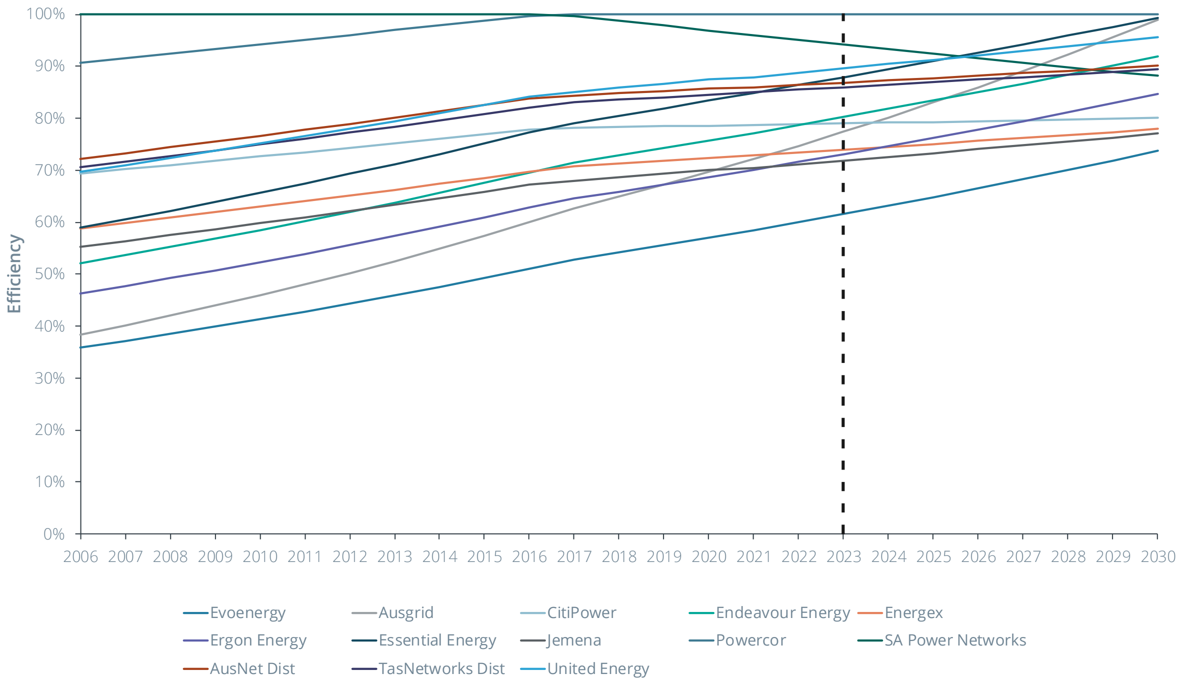
Source: Frontier Economics.

Figure 95: Efficiency predictions under constant efficiency data generation – LSE-AJTT-NZ-TL



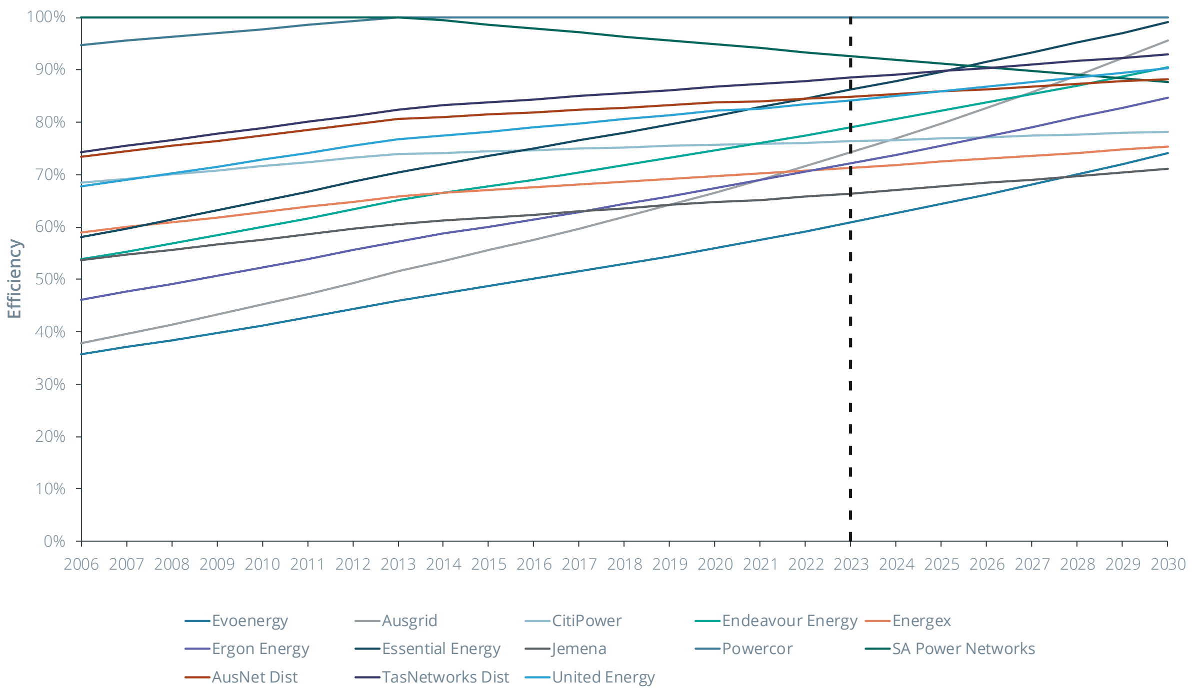
Source: Frontier Economics.

Figure 96: Efficiency predictions under constant efficiency data generation – LSE-AJTT-NZ-GTC-TL



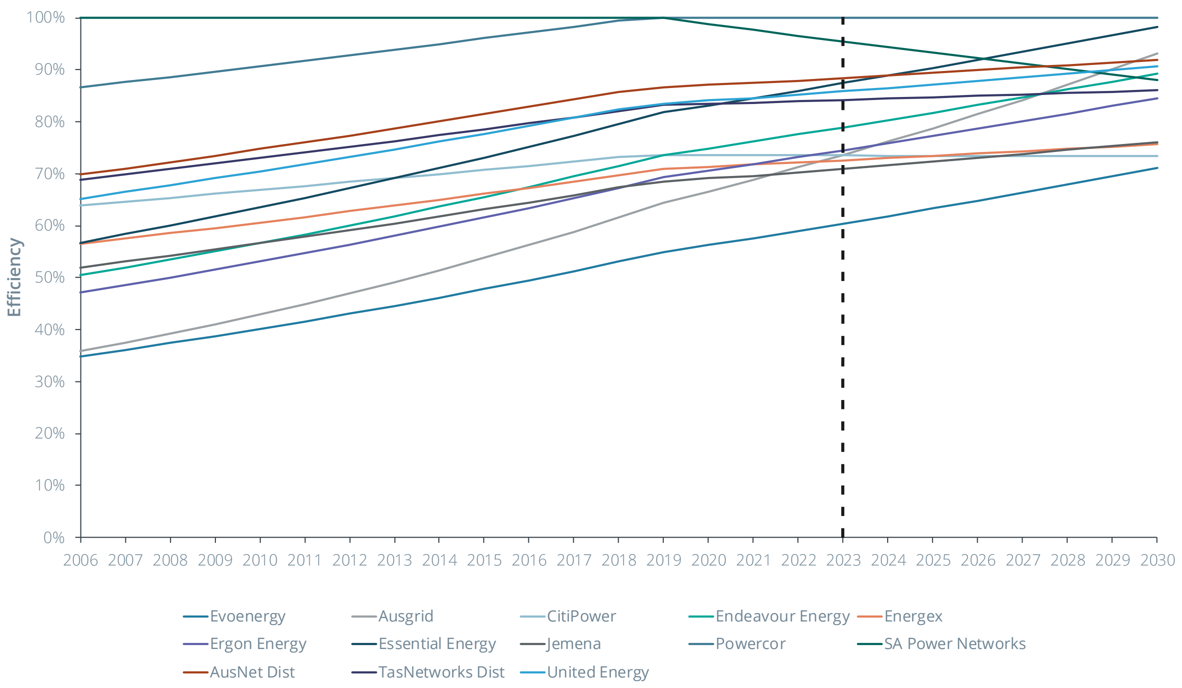
Source: Frontier Economics.

Figure 97: Efficiency predictions under constant efficiency data generation – LSE-AJTT-ONT-TL



Source: Frontier Economics.

Figure 98: Efficiency predictions under constant efficiency data generation – LSE-AJTT-ONT-GTC-TL



Source: Frontier Economics.

Frontier Economics

Brisbane | Melbourne | Singapore | Sydney

Frontier Economics Pty Ltd
395 Collins Street Melbourne Victoria 3000

Tel: +61 3 9620 4488
www.frontier-economics.com.au

ACN: 087 553 124 ABN: 13 087 553 124