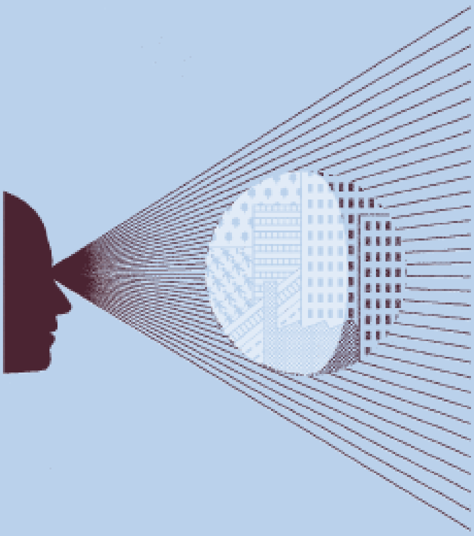
Recommendations on cost assessment approaches for RIIO-ED1

An independent submission by Oxera to Ofgem

In association with Distinguished Professor Subal Kumbhakar

February 2013



Executive summary

For RIIO-ED1, Ofgem has stated its intention to use a 'toolkit' approach to determine the relative efficiency of the GB electricity DNOs. However, Oxera understands that, in a recent Cost Assessment Working Group (CAWG) meeting, Ofgem stated that, unless there is a rationale for opting for approaches that could control for company-specific factors within modelling, it is currently minded to focus on pooled ordinary least squares (OLS) as the econometric method to estimate the DNOs' efficiency levels.

This report shows that there are few alternative approaches that are superior to pooled OLS and that Ofgem could consider these as part of its cost assessment toolkit for RIIO-ED1 as they provide more accurate assessments of the inefficiency levels. Indeed, the proposed use of pooled OLS could be replaced with more robust models that have been developed in the efficiency literature over the last 40 years, many of which have been used by regulators in various jurisdictions.

The report also demonstrates, both theoretically and empirically, that pooled OLS leads to incorrect conclusions when there are company-specific effects, as it overestimates inefficiency. With the current dataset, this overestimation is of the order of doubling the industry average inefficiency gap. Since inefficiency is a residual measure, care must be taken in modelling the cost drivers and choosing appropriate model and estimation technique so that the efficiency frontier is estimated consistently and accurately. It is shown that OLS estimates can be biased and thereby result in incorrect estimates of inefficiency (no matter how inefficiency is calculated), because the residuals on which inefficiency measures are based are incorrect. Pooled OLS can produce biased and inconsistent parameter estimates if inefficiency changes over time, if there are company-specific effects, and if inefficiency depends on some exogenous variables (observed or unobserved), such that the inefficiency estimated from the model cannot be relied on. These contexts all seem likely and thus it seems likely that pooled OLS cannot be relied upon. As such, Oxera would recommend examining various panel data[1] and stochastic frontier analysis (SFA) modelling approaches in comparison with the pooled OLS model. In all the models, the OLS model is a special case and therefore the OLS specification can be statistically tested.

These alternative models provide more robust estimates of the DNOs' relative efficiency levels and hence improve Ofgem's efficiency analysis. The approaches that provide the most accurate assessments could also be used to establish the extent to which a simpler approach requires company-specific adjustments in order to account for modelling errors and company-specific effects.

As Ofgem's preferred modelling approach—pooled OLS—overestimates inefficiency, it is not possible to amend the OLS estimates objectively in order to derive the 'true' inefficiency gap. Instead, to derive more reasonable inefficiency estimates, ad hoc adjustments are required. Ofgem's approach has previously used an upper-quartile or upper-third benchmark to determine the gap, depending on the range of the estimated gaps. However, the upper-quartile (and upper-third) correction is ad hoc as it assumes that the same level of noise is present in the data across all the companies, thereby potentially under- or over-correcting for some DNOs. That is, the inefficiency of some companies will still be overestimated, while that of others will be underestimated. This is demonstrated empirically with the current dataset: while, at the industry level, the upper-quartile benchmark provides a reasonable correction, at the DNO level, in most instances, the inefficiency gap is either under- or

[1] That is, data for a set of observations over time.

overestimated. In particular, on the current dataset, eight DNOs' efficiency gaps are over-corrected for noise with a maximum difference of +7% and six DNOs' efficiency gaps are under-corrected for noise with a maximum difference of –13%, with two DNOs' differences in the gaps less than ±1%. At the total expenditure (TOTEX) level, this subjective adjustment could result in a considerable difference in monetary terms, and, as such, could constitute a significant regulatory risk.

In addition, Ofgem's current approach necessitates regulatory judgement to normalise for DNO-specific differences. As the sizes of the cost adjustments are likely to be difficult to quantify, it represents an additional regulatory risk and may further reduce the accuracy of the cost assessment approach. In contrast, there is a wealth of literature on efficiency estimation within the panel framework that can lessen this need for ad hoc assumptions and unnecessary regulatory judgement.

Based on theoretical foundations and empirical analysis (which confirms the theoretical insights), this report provides a number of findings and recommendations. These recommendations will result in a more robust estimation of the DNOs' relative efficiency levels than the use of pooled OLS alone, or could be used to provide a more robust basis for establishing ad hoc adjustments for each DNO.

– It is shown that pooled OLS estimates (with and without year dummies) can be biased and thereby result in incorrect estimates of inefficiency (no matter how inefficiency is calculated), because the residuals on which inefficiency measures are based are incorrect.

– SFA and data envelopment analysis (DEA) have regulatory precedent in the UK and across Europe. To increase the robustness of the results, Oxera recommends that a number of approaches are examined, including SFA and DEA. In addition, all the models estimated in this report can be implemented using publicly available software tools and are well established in the academic literature.

– The results from the different approaches should be compared and contrasted and, based on an understanding of the approaches, some consensus could be reached in order to identify a more robust range for the estimated inefficiencies.

– SFA and panel data models allow for direct interpretation of the residuals. In all the SFA models, the statistical significance of inefficiency can be tested. This is not the case with OLS (or corrected OLS, COLS). Indeed, further extensions in an SFA panel setting allow for explicit interpretation of the results in terms of uncontrollable company-specific effects, noise in data/modelling errors, persistent inefficiency and transient inefficiency. Such a decomposition and interpretation is currently not possible using other approaches.

– The alternative models examined in this report have been shown to be both valid and practical for the current dataset in terms of statistical robustness and economic interpretation of the results.

– While Oxera would recommend examining panel data modelling approaches, it also appears to be the case that the 'fixed-effects' panel model is not appropriate in this dataset, as some of the cost drivers do not change much over time. A 'random-effects' approach would therefore be more appropriate. This is also empirically determined on the current dataset using a statistical test.

– In this instance, with the dataset examined in this report, the ad hoc adjustment of an upper quartile is close, at an industry-wide level, to making an appropriate adjustment for errors. This could be coincidental and would need to be checked in other circumstances.

– However, while the average gap is similar to that of other models following this adjustment at the industry level, at an individual DNO level there is considerable variation in the estimated gaps—in some cases the adjustment is too much and in others it is too little. At the TOTEX level, such differences could be significant in monetary terms.

Finally, in this report it has only been possible to examine historical costs and cost drivers empirically. Ofgem may, however, place weight on efficiency assessments made using companies' forecast data, as it did in RIIO-GD1. Such forecast data (over a period of eight years) is likely to show far more variation over time than actual historical data, and, in such a context, separating efficiency that does not change over time from transient inefficiency, and from errors and company-specific effects, may be more important than in the current context.

Contents

List of tables

List of figures

List of boxes

1 Introduction

Regulators use cost benchmarking widely to set cost allowances, and hence cost-reduction targets for companies to achieve, during price reviews or to monitor their performance. Companies also use it to determine internal efficiency challenges for business planning purposes (or for presenting well-justified business plans to the regulator). For example, in DPCR5, Ofgem applied econometric modelling on a panel dataset (ie, data on companies over time) to estimate the efficient OPEX levels of the DNOs.

The pooled ordinary least squares (OLS) model adopted by Ofgem in DPCR5 ignores the panel structure of the dataset and applies an OLS technique on the pooled dataset with year dummies to capture the movement in the average cost over time. In the ongoing RIIO price reviews, in particular RIIO-GD1 and RIIO-ED1, Ofgem adopted or has proposed to adopt a similar approach to estimate the efficient cost levels at the total expenditure (TOTEX) level, as well as at the disaggregate levels: OPEX; capital expenditure (CAPEX), and replacement expenditure (REPEX).

As part of DPCR5, Ofgem also considered a panel modelling approach that controls for DNO-specific effects within the model, noting that the technique would enable ease of replication.[2] Ofgem offered two reasons for its preference of a pooled OLS over a panel modelling approach: a pooled OLS approach with regulator–DNO dialogue to adjust for the DNO-specific differences would provide 'greater transparency';[3] and the robustness of the results from the DNO-specific fixed-effects model could require a longer time series than the four years of data that was considered in the analysis.

The validity of the first reason is not clear, as Ofgem undertook 27 DNO-specific cost adjustments (17 cost exclusions and ten normalisations) on a case-by-case basis to take into account factors outside the DNOs' control that have an impact on their cost performance.[4] These adjustments are likely to involve a degree of subjectivity. In particular, Ofwat argues that the size of these cost adjustments is often difficult to quantify, and that an attempt to correct modelled cost by normalising for company-specific differences might introduce noise to the data and reduce the accuracy of the assessment instead of improving it.[5] For this reason, Ofwat models unadjusted costs and applies the adjustments post-modelling when setting the overall cost-reduction target for the companies.

While Ofgem's second reason is theoretically valid, unless the firm effects are controlled for, the parameters of interest (ie, the regression estimates, and thus the estimated efficiency levels) might be biased. In contrast, a panel modelling approach has the potential to remove some of this regulatory judgement as the DNO-specific adjustments can be controlled for by a DNO-specific-effects model, and the validity of these approaches can be tested for. Finally, it is not clear that Ofgem has thoroughly investigated such alternatives and the feasibility of panel modelling approaches. This report provides such an investigation, with the aim of adding to the current debate in order to mitigate the risk of superior efficiency estimation approaches being overlooked.

[2] Ofgem (2009), 'Electricity Distribution Price Control Review Initial Proposals—Allowed revenue—cost assessment', August 3rd, Appendix 9, para 1.10.

[3] Ibid.

[4] Ibid., pp. 62–7.

[5] See slide 3 of http://www.ofwat.gov.uk/legacy/aptrix/ofwat/publish.nsf/AttachmentsByTitle/ms_part2110707.pdf/$FILE/ms_part2110707.pdf; accessed on February 5th 2013.

Indeed, such approaches have been used in other regulatory contexts. For example, a brief survey of the cost benchmarking approaches adopted by other UK regulators in their most recent price control reviews, or which they have proposed to use for an ongoing review, indicates that they rely heavily on panel data modelling techniques. For example, the ORR[6] and Ofcom[7] consider a panel stochastic frontier analysis (SFA) approach as part of the cost assessment toolkit. In addition, Ofwat is considering panel data modelling techniques, including the random-effects model and panel SFA, to assess the TOTEX efficiency of the England and Wales water and sewerage companies for the ongoing price review.[8]

1.1 Ongoing cost benchmarking work for RIIO-ED1

Oxera understands that Frontier Economics has been developing TOTEX benchmarking models for the electricity DNOs. Ofgem has indicated that it would adopt the models as they are, or modify them, or reject them altogether and develop alternative ones.[9] Frontier Economics' preferred model uses data over five years (2006/07–2010/11), with TOTEX as the cost measure, and number of customers, peak capacity, population density and national wage index as explanatory variables. A time trend is included in the model to control for movement in costs over time, as a proxy to measure the technological change in the industry over the period. The estimation is undertaken using a statistical model (a random effects model) that attempts to estimate a company-specific component which is assumed to be invariant over time. This company-specific component is taken to be the measure of efficiency. As such, the approach is similar to that considered by Ofgem in DPCR5.

1.2 Objective of this report and remit

ENWL commissioned Oxera to examine alternative and potentially more robust econometric modelling approaches to estimate the DNOs' efficiency levels. The material used as the input for this report was limited to the presentation slides from Frontier Economics and the dataset that it used in its analysis. Both were provided to Oxera by ENWL. Oxera was not involved in the Cost Assessment Working Group (CAWG) meetings and any discussions the DNOs may have had with Frontier Economics. As such, there may be issues identified by Oxera in this report that Frontier Economics or the DNOs have already considered, and on which both parties might already have reached an agreement; equally, there may be further issues not examined in this report that have previously been raised.

As such, this report does not examine many important issues such as translating the results into cost allowances, the functional form of the model, the definition of the modelled costs, the variables used in the model, approaches outside of standard econometric approaches (such as data envelopment analysis, DEA), etc.

Within the constraints of the remit, including the materials examined, this report examines various econometric models that Ofgem could consider as part of its cost assessment exercise for RIIO-ED1, including consideration of:

– the academic basis for these models;
– the models' advantages and disadvantages;

[6] See ORR (2011), 'Establishing Network Rail's efficient expenditure', July, pp. 28–30. For the price control review for the period 2013–18, the ORR is proposing to use advanced SFA models to assess the efficiency of Network Rail, based on recommendations made in Oxera (2009), 'Recommendations on how to model efficiency for future price reviews', November.

[7] For more information, see NERA (2008), 'The comparative efficiency of BT Openreach', a report for Ofcom, March.

[8] See CEPA (2013), 'PR14 Cost Assessment', a report for Ofwat, January.

[9] See Ofgem (2012), 'Cost Assessment Working Group', Meeting 6, presentation, July 31st, available at http://www.ofgem.gov.uk/Networks/ElecDist/PriceCntrls/riio-ed1/working-groups/Documents1/Ofgem_presentation_CAWG_31072012.pdf; accessed February 5th 2013.

– the feasibility of applying these models using the dataset employed by Frontier Economics in its analysis and their statistical validity.[10]

1.3 Structure of report

The remainder of this report is structured as follows.

– Section 2 provides background to cost efficiency benchmarking in a regulatory context and provides a number of recommendations.

– Section 3 describes the panel models considered in the report, including their advantages and disadvantages.

– Section 4 discusses the results from these models, including their statistical validity using the current DNO dataset.

– Section 5 concludes.

– Appendix 1 describes some cost benchmarking approaches adopted by UK regulators.

[10] The slides and dataset from Frontier Economics were received from ENWL.

2 Background to cost benchmarking in a regulatory context

Regulators use cost benchmarking widely to set cost allowances during price reviews or to monitor companies' performance, and by companies to determine internal efficiency challenges. This section examines cost benchmarking as used in a regulatory context, and looks at some specific issues with regard to the dataset available for RIIO-ED1.

2.1 Objectives of cost benchmarking (in a regulatory context)

The overall aim of a benchmarking exercise is to establish the scope for efficiency improvements that a company can achieve, and where those efficiencies can be achieved. For regulatory purposes, it is the former issue that is the primary focus of the analysis.

2.1.1 Catch-up and frontier shift

Theoretically, the scope for efficiency improvements has two components:

– **catch-up**, which provides an estimate of the potential to catch up to current best practice. Estimates of the catch-up potential for a company are often based on estimates of current relative efficiency—ie, the current gap to the best-performing companies;

– **frontier shift**, which provides an estimate of the likely productivity improvements that the assessed company can make by adopting new technologies and working practices, above and beyond any cost reductions resulting from the company improving its relative efficiency. The frontier-shift target is set for each company in the industry in addition to any catch-up assumption.

Some assessment approaches allow for both catch-up and frontier shift to be estimated within the same methodological framework; alternatively, these two components have been estimated separately using a mixture of approaches. Within a panel modelling framework it is possible to derive estimates for both components.

Recommendation 1: to enable an examination of *both* catch-up and frontier shift, Oxera would recommend the use of a panel dataset, as Ofgem does in its approach. As the objective of an efficiency analysis is not only to estimate the efficiency gap of the companies relative to the current best practice (ie, technology), but also to determine how the technology is changing over time, it is preferable to estimate both factors simultaneously from the same model. This could reduce the reliance on indirect comparisons (such as the productivity performance of other sectors in the economy, which is often used as a basis for identifying a frontier shift). However, SFA using a panel dataset can separate these two elements, while Ofgem's current model, pooled OLS, or standard fixed- or random-effects panel models cannot (see section 3).

2.1.2 Estimating efficiency

A key objective is to be able to estimate relative inefficiency. Before examining this issue, some background in terms of data availability and DNO-specific effects are introduced; methods to estimating relative efficiency and issues surrounding this are discussed in section 2.4.

2.2 Data used for the analysis: the size of the dataset and accuracy of the modelling

Ofgem uses cost and cost driver data, provided by the DNOs in their annual Regulatory Reporting Pack (RRPs), as well as unit cost information from technical reports and data in the RRPs for the calculation of variables, such as the modern equivalent asset value (MEAV).

With 14 DNOs and six independent groups, the cross-sectional information is relatively limited. However, the dataset can be readily increased by including data on the companies over time, forming a panel dataset. The increased accuracy from such an approach has been demonstrated in the literature in a regulatory context.[11]

Data inconsistencies across years and between DNOs have meant that the current Frontier Economics analysis is based on only five years of data. (Frontier Economics has furthermore indicated that there are issues with the first two years of cost data, essentially limiting the panel dataset to the last three years.) With 14 DNOs and six independent groups, this provides 70 observations (14 DNOs for five years, or 42 observations if only three years are used), and 30 independent observations (six ownership groups for five years, or 18 observations if only three years are used). This limits the number of explanatory factors that can be included in the model. To mitigate the data limitations, in DPCR5 Ofgem used a composite scale variable (CSV) where there is estimated to be more than one cost driver. The small number of independent observations also means that the results can be skewed by one or two observations, which can further limit the robustness of the model.

Recommendation 2: to increase the number of observations for modelling, Oxera would recommend the use of a panel dataset, as Ofgem uses in its approach. Extending the number of years of data covered in the dataset helps to increase the size of the panel dataset, improving the robustness of the models and the estimated efficiency levels of the DNOs (and limits the need to use a CSV). Extending the dataset raises the possibility of the cost relationships changing over time,[12] and thus the models and the results changing, which is likely if forecast data is used in the analysis. The potentially more robust panel models, some of which are discussed in section 3, may be practical and useful in such instances (ie, where forecast data is assessed either separately or together with historical actual data). Oxera would also recommend that emphasis be placed on industry knowledge and the alignment of the models with this industry knowledge.

Panel data analysis is analysed in more depth below, and, in particular, contrasted with pooled OLS.

2.3 Company-specific effects: pooled OLS versus panel data modelling

Econometric modelling can take into account only a limited number of external factors[13] that result in costs differing between DNOs. As a result, other factors, such as company characteristics, that are not suitably accounted for in the regression modelling also need to be taken into consideration. There are two possible approaches for taking into account such company-specific effects:

[11] See, for example, Kumbhakar, S. and Horncastle, A. (2010), 'Improving the Econometric Precision of Regulatory Models', *Journal of Regulatory Economics*, **38**:2, October, pp. 144–66.

[12] Such changes can and should be tested for.

[13] Especially in this case, where the available cross-section of the dataset is relatively small (14 DNOs and 6 ownership groups).

– company-specific adjustments, whereby the impact of external DNO differences is quantified separately and taken into account either before or after the econometric modelling;
– using panel data to take account of company-specific effects (random or fixed effects).

As **Pooled OLS** pools the cross-sectional data over time into one dataset and models the data in the standard cross-sectional OLS way, it ignores the time and company dimension of the dataset by treating each observation as a different DNO. This **can result in overestimating the inefficiency levels of the DNOs as it cannot separate inefficiency from firm effects and noise**.

In contrast, panel data modelling allows company-specific factors that cannot be observed or suitably measured, such differences in topography, condition of assets, or variables that change over time but not across entities (eg, energy policies, regulatory changes), to be controlled for. That is, panel data can account for individual dissimilarity (heterogeneity).

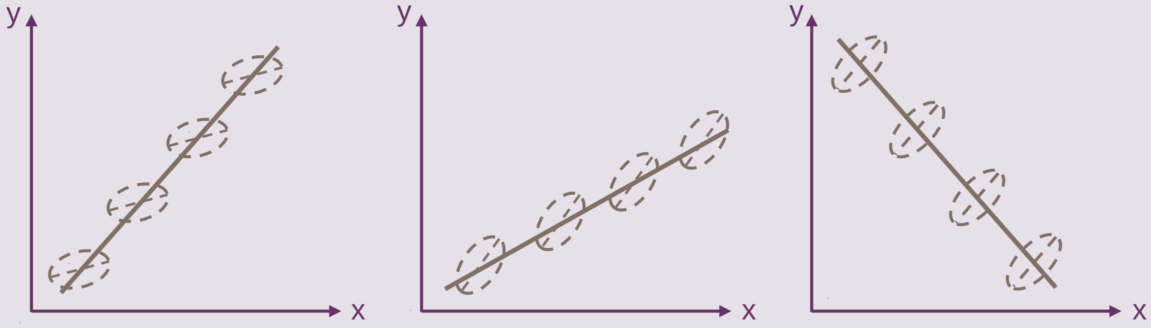
The limitations of the pooled OLS approach compared with a panel approach are illustrated below.

Box 2.1 Comparison of pooled OLS and panel data modelling

The figures below illustrate potential cases with a cross-section of four companies observed over a period of time. In each case the relationship between costs, y, and output, x, is positive, but each company is different in some way and these differences are captured by company-specific intercepts (ie, the fixed cost of being in business depending on its operating conditions), but a common slope (ie, variable cost). The differences in the intercepts can also be due to inefficiency. As a result, when pooled OLS is applied to the data, which estimates a common intercept and slope across the companies, the estimated relationship between cost driver x and cost y is very different from the true relationship, and may therefore be meaningless. While pooled OLS may be valid, it is critical that this is tested for in the dataset.

In the first figure below (to the left), the estimated slope using pooled OLS is steeper than the true relationship; in the second (in the middle) the estimated slope is flatter than the true relationship; and in the third (to the right), the estimated slope is actually negative. In an efficiency context estimating the true relationship is critical, as the estimated residuals, and the resultant inefficiency levels, will also be incorrect if the cost function is incorrectly estimated.

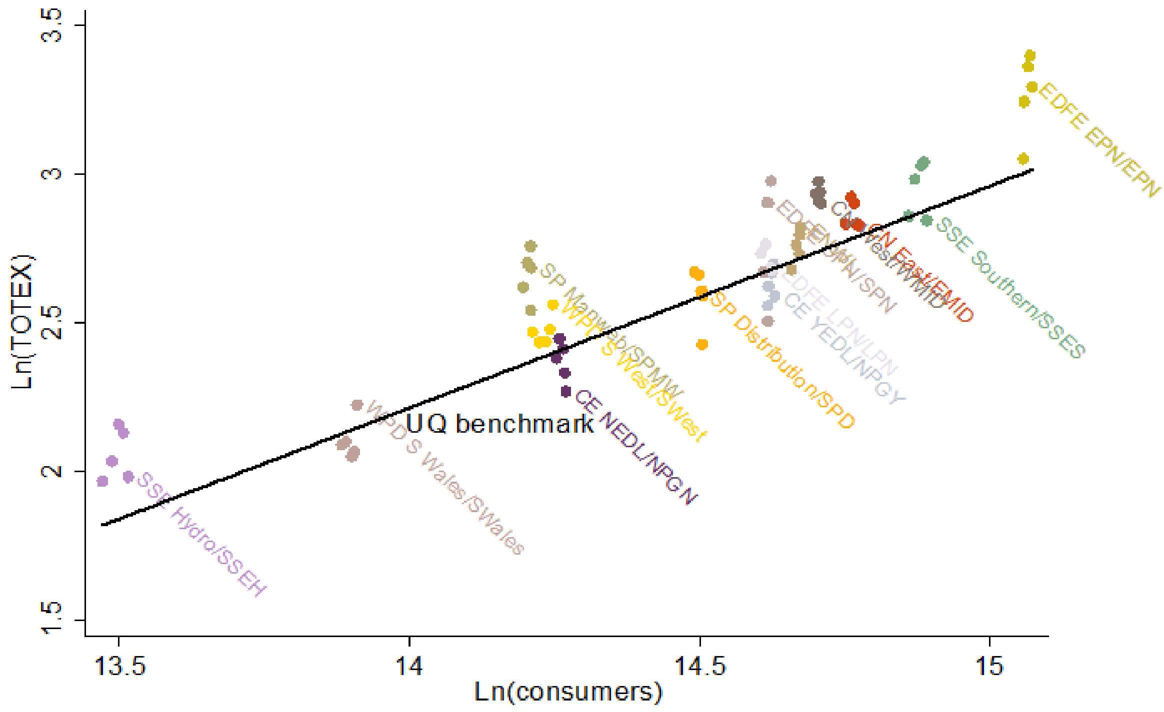
Graphical illustration of pooled OLS and panel data modelling



Note: These figures are intended as an illustration only.

Figure 2.1 examines the above illustrations using the actual DNO data over the period 2006/07 to 2010/11.

Figure 2.1 Illustration of pooled OLS and panel modelling using DNO data



Note: The above figure illustrates the spread of data on TOTEX and the number of consumers of the companies over the five-year period, with the UQ benchmark representing the OLS regression line corrected to the upper quartile of the efficiency levels.
Source: Oxera analysis based on data from Frontier Economics provided by ENWL.

As shown in Figure 2.1, and consistent with the discussion in Box 2.1, the TOTEX differences between DNOs with similar consumer numbers can be large, but these are closely clustered for each DNO over the period. For example, DNOs CE NEDL, WPD S West and SP Manweb appear to have similar consumer numbers over the same time period, although CE NEDL’s TOTEX is lower than the other two. A simplistic view of cost assessment would suggest that WPD S West and SP Manweb are far more inefficient than CE NEDL; while, directionally, this may be correct, much of the relative gap may be explained by other factors that are different between them as they each serve different areas. By making use of the panel data and controlling for company-specific effects, the model has the potential to estimate the relative efficiency of the DNOs more robustly.

In terms of panel data models, from the spread of the data over the period, it appears that a fixed-effects model may not be appropriate as it cannot separate between time-invariant explanatory variables (in the case of the Frontier Economics model, for example, population density) from time-invariant company effects (and, thus, the estimated efficiency levels). There appears to be a consistent movement in average cost over time, which could also cause issues for pooled OLS. (Simply adding time dummies to pooled OLS would still result in incorrect estimation of the regression model, unless the company-specific effects and noise in the data are accurately accounted for.)

It is clear from the above discussion that these company-specific effects need to be taken into account when estimating inefficiency.

Recommendation 3: *company-specific adjustments* are a valid approach when taking into account variations in costs driven by exogenous factors (eg, company-specific effects) that are not accounted for in the regression modelling. These factors should be estimated with respect to the regression models used. However, it is unlikely that all company-specific effects (and certainly beneficial ones) will be captured using this approach, with the result that some company-specific effects are likely to remain. For example, Ofwat and its academic advisers have noted the uncertainty involved in these cost adjustments, which are likely to reduce the accuracy of the cost assessment method..[14] Given this, panel data analysis should be used to take account of company-specific effects. However, it also appears to be the case that the fixed-effects panel model is not appropriate in this dataset because some of the cost drivers do not change much over time, so a random-effects approach is more appropriate.

In addition, as demonstrated above, in the presence of fixed firm and time effects, alongside time-varying inefficiency, OLS parameters are likely to be inconsistent, and therefore the OLS residuals are unreliable for estimating inefficiency via COLS. While pooled OLS estimates will be consistent in the presence of random firm effects, the estimated standard errors will be incorrect and they cannot be used for hypothesis testing (and thus interpreting the fit of the model to the dataset). If inefficiency is time-varying, even if firm effects are random, the estimated OLS coefficients will be biased and the OLS residuals cannot be relied on for estimating inefficiency.

Recommendation 4: pooled OLS can be misleading when there are company-specific effects or when inefficiency varies over time.[15] As such, Oxera would recommend examining panel data modelling approaches.

2.4 Estimating inefficiency and interpreting residuals

The discussion so far has focused on the econometric estimation of the cost function. From this, an estimate of relative inefficiency needs to be derived, which can be done using several analytical approaches. The brief survey presented in Appendix 1 indicates that the techniques commonly used by UK regulators are SFA and DEA.

Recommendation 5: SFA and DEA have regulatory precedent in the UK and across Europe. To increase the robustness of the results, Oxera recommends that a number of approaches are examined, including SFA and DEA. The results from the different approaches should be compared and contrasted, and, based on an understanding of the approaches, consensus could be sought in order to identify a more robust range for the estimated inefficiencies.

Given that this report focuses on econometric approaches, this section provides a brief technical description of COLS and SFA. Both approaches use the residuals from the econometric model to estimate inefficiency, so their interpretation and validity as a basis for estimating inefficiency is critical.

2.4.1 Assuming there are no modelling or data errors: COLS

An econometric approach based on a simple regression model (OLS[16]), COLS is a frontier-based approach, in that it measures efficiency by reference to an efficiency frontier. The

[14] See slide 3 of Stewart, M.(2007), 'Development of Econometric Models for Capital Maintenance Relative Efficiency Assessment: Part 2', presentation at Capital Maintenance Relative Efficiency Modelling Workshop, Ofwat , July 11th, 2007http://www.ofwat.gov.uk/legacy/aptrix/ofwat/publish.nsf/AttachmentsByTitle/ms_part2110707.pdf/$FILE/ms_part2110707.pdf; accessed on February 5th 2013.

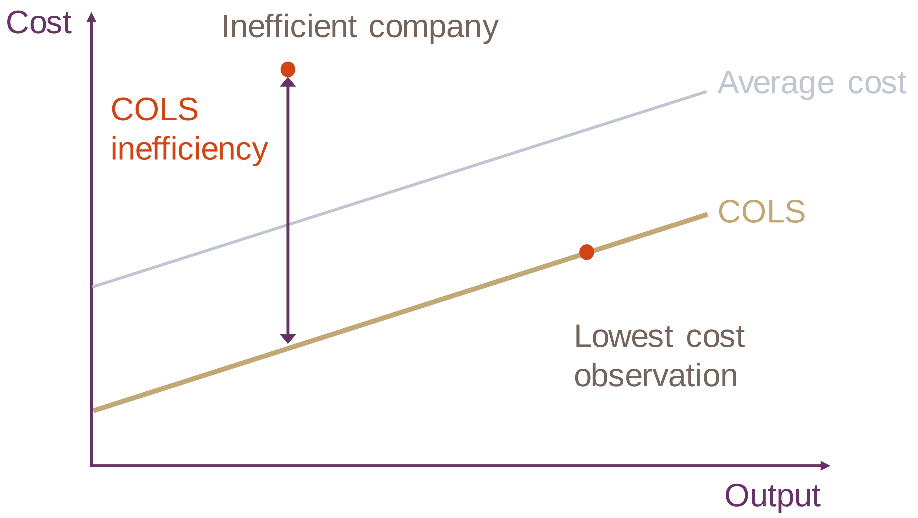
[15] In the presence of fixed firm and time effects, as well as time-varying inefficiency, parameters estimated using OLS are likely to be inconsistent.

[16] OLS estimates a linear relationship between a dependent variable, here TOTEX, and a set of cost drivers, here number of consumers, peak capacity and population density. It does this by minimising the distance between each observation and the estimated regression line (technically, minimises the sum of squared residuals).

frontier is derived by shifting the OLS line of best fit from the estimated cost function of the industry, so that, instead of representing the average cost of the industry, the line represents the efficiency frontier (ie, companies operating using the least cost). This shift can be based on the maximum negative residual of the regression model, resulting in the pure COLS frontier. If the residuals are not consistently estimated, the resulting frontier might be incorrect. For example, the presence of outliers can affect the frontier significantly. Similarly, good or bad weather in a particular year might be confounded with inefficiency. This suggests that the 'constructed frontier' is likely to be different from the 'true frontier', and there is no way of testing this econometrically under COLS.

Figure 2.2 provides an illustration of the COLS approach.

Figure 2.2 COLS frontier and efficiency



Source: Oxera.

Inaccuracies and other statistical noise in the model mean that the gap between the most efficient DNO and the other companies is not due entirely to inefficiency. Since a company can be estimated to be efficient simply because its inefficiency is confounded with noise, it is essential to separate noise from inefficiency. There are several approaches that deal with this uncertainty. Before examining two main approaches used by regulators in the econometric framework to deal with uncertainty, a theoretical examination of why COLS may not be appropriate for efficiency assessments is discussed in Box 2.2.

Box 2.2 Theoretical examination of why COLS may not appropriate for efficiency assessments

Consider a cross-sectional setting where the relationship between cost y and cost drivers x is estimated using OLS regression. The model can be written as in Equation 1:

$$Y_i = a + bX_i + v_i \qquad \textbf{Equation 1}$$

In OLS, the term v_i denotes the noise or the error term which is assumed to be normally distributed (with zero mean and constant variance). If efficiency is estimated from the model as $(\hat{v}_i - min\,(\hat{v}_i))$ as Ofgem currently does, there is a 'philosophical problem' in that inefficiency is estimated from noise.

If v_i is viewed as inefficiency then, by definition, it is skewed and non-negative (as a firm is either efficient and thus on the efficiency frontier, or inefficient) and the expected value of the term denoted by $E(\hat{v}_i)$ will not be zero. To circumvent this issue, Equation 1 can be rewritten as

$$Y_i = a + bX_i + E(v_i) + \big(v_i - E(v_i)\big) \qquad \textbf{Equation 2}$$

If the error term is expressed as $q_i = \left(v_i - E(v_i)\right)$, it has zero mean by construction. If $E(\widehat{v_i})$ is a constant it makes the OLS intercept $\widehat{a}$ biased. However, if inefficiency is estimated using $(\widehat{q_i} - min\,(\widehat{q_i})) = (\widehat{v_i} - min\,(\widehat{v_i}))$, inefficiency, being a relative measure, can still be properly assessed.

This view of inefficiency using COLS goes back to the early 1970s wherein the assumption that the mean of inefficiency is a constant had to be made. However, almost every model that has been recently developed does not assume this and explains inefficiency as a function of some exogenous variables (or determinants of inefficiency), say Z. In such a case $E(v_i)$ will be a function of Z_i, say $m(Z_i)$. This means that Equation 2 becomes:

$$Y_i = a + bX_i + m(Z_i) + \left(v_i - E(v_i)\right) \qquad \textbf{Equation 3}$$

Since $m(Z_i)$ is not a constant, it cannot be absorbed by the intercept. So, if one runs the OLS regression Y on X, the model will suffer from omitted variable bias due to excluding $m(Z_i)$. If Z is correlated with X (which is most likely the case in any economic problem), the OLS estimates will be inconsistent.

However, the problem with using OLS does not end here. When inefficiency is estimated using the OLS residuals as $(\widehat{q_i} - min\,(\widehat{q_i})) = (\widehat{v_i} - E(\widehat{v_i})) - min(\widehat{v_i} - E(\widehat{v_i}))\,) = \widehat{v_i} - min(\widehat{v_i}) - E(\widehat{v_i}) - min(E\,(\widehat{v_i}))$, the adjusted OLS residuals cannot be interpreted as relative inefficiency since $E(\widehat{v_i})$ and $min(E\,(\widehat{v_i}))$ are not the same.

So there are two major problems here: the OLS estimates (parameters) are inconsistent, and the COLS formula does not give a measure of relative inefficiency.

Furthermore, the estimated standard errors will be inconsistent if variance of v_i is not constant. Thus, there is no basis to rely on COLS.

When a panel dataset is available, the data is pooled and OLS is run on the pooled data, there are additional problems. If it is assumed that there are no firm effects (fixed or random), the problem noted earlier will apply if v_{it} is interpreted as inefficiency. If $E(v_{it})$ is a constant (which is highly unlikely), and relative efficiency is measured from OLS residual e_{it} minus the minimum value of e_{it} (over i and t), it will represent inefficiency relative to the best-performing firm over all the years. This is likely to overestimate inefficiency. If, instead, $E(v_{it})$ is not a constant but varies over time, for example, the inconsistency problem recurs due to an omitted variable (ie, due to not including $E(v_{it})$ in the regression).

SFA is a possible solution to the above problem.

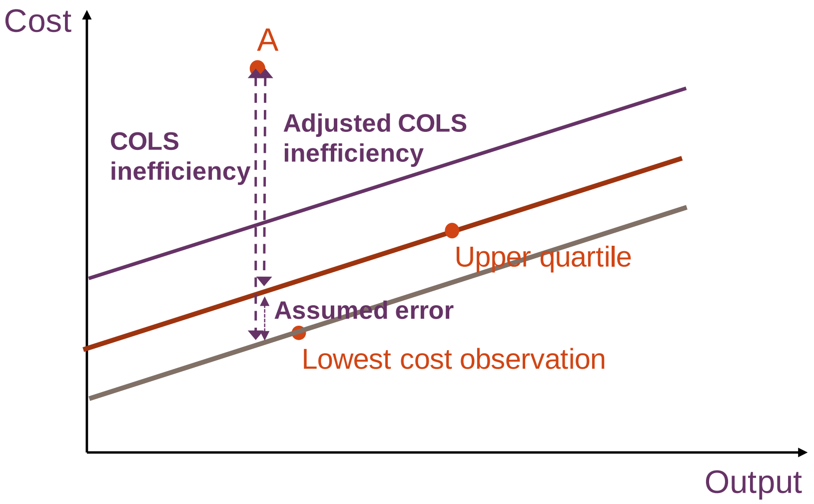
2.4.2 Accounting for noise

Accounting for noise: ad hoc adjustments to COLS

One of the main weaknesses of the COLS approach is that it assumes that any difference between a company’s observed costs and the regression line (ie, the residual) represents inefficiency. It does not account for any stochastic error or noise in the model, such as measurement error, which affects the size of the residual. COLS can therefore overestimate inefficiency when the error component of the residual is large. Also, the model is deterministic, in that it does not allow for statistical testing (eg, confidence intervals) of the resultant efficiency estimates.

Alternatively, an ad hoc adjustment of the COLS frontier may be applied. For example, Ofgem has historically used an upper-quartile or an upper-third benchmark, depending on the range of the estimated efficiency levels. That is, under the upper-third benchmark, the assumption is that the difference between the frontier company and the upper-quartile company represents the noise in the estimated inefficiency for every other company. However, the upper-quartile correction is ad hoc and assumes that the same level of noise is present in the data across the companies, thereby potentially under- or over-correcting for some. This approach is illustrated in Figure 2.3 below.

Figure 2.3 Estimating inefficiency using COLS with an upper quartile benchmark



Source: Oxera.

Accounting for noise: stochastic frontier analysis

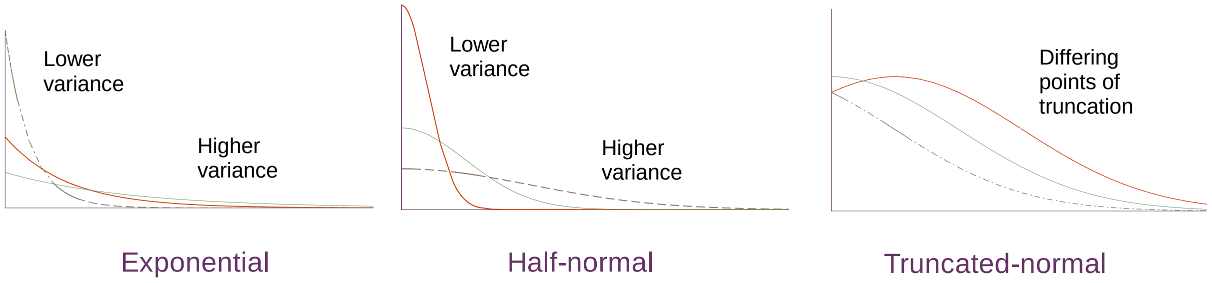
A potentially more appropriate and less ad hoc way to adjust for uncertainty is to use SFA. As well as being used in other regulatory jurisdictions, SFA has a long academic history, having first been developed in 1977.[17] There are thousands of theoretical and applied papers on the approach and several applied in different jurisdictions in a regulatory context.

As explained above, when companies are not fully efficient and the objective is not only to estimate the parameters but also to predict cost inefficiency, OLS is not the most appropriate approach. OLS residuals contain both noise and inefficiency, and there is no way of separating inefficiency from noise, given the OLS residuals. The use of an upper-quartile benchmark is just an assumption. Furthermore, OLS estimates (and therefore the residuals) might be inconsistent if firm effects are ignored and a pooled model is used.

SFA is an econometric method that estimates the parameters of a cost function and the inefficiency for each observation. The separation of inefficiency from the residuals is achieved with the help of the distributional assumption made on the noise and the inefficiency components. The error component of the residual is assumed to be symmetric (noise can have both a positive and a negative effect), while the inefficiency component is asymmetric, in that a company is either efficient or inefficient (it cannot be better than the efficiency frontier). Specifically, the error component of the residual is assumed to be normally distributed (as it is in the standard OLS approach at least for hypothesis testing), while the inefficiency component is assumed to be half-normal, truncated or exponentially distributed. This is the assumption required to separate noise from inefficiency (as illustrated in Figure 2.4 below).

[17] For a more detailed discussion on SFA, see Kumbhakar, S.C. and Knox Lovell, C.A. (2000), *Stochastic Frontier Analysis*, Cambridge University Press.

Figure 2.4 Distributional assumptions for SFA, with different variances or points of truncation (in the case of the truncated-normal)



After more than 20 years of regulation and comparative efficiency-based cost allowances, DNOs would now be expected to be converging towards the frontier, such as in the first two distributional assumptions presented above (where the variance of the efficiency gap is small). If, despite such regulatory incentives, there are reasons to expect that the majority of DNOs are instead clustering slightly away from best practice in the industry, the truncated-normal distribution (illustrated on the right) may be deemed more appropriate. Alternatively, in the half-normal case, the variance could be high (and it is possible to test whether a truncated-normal or a half-normal is more appropriate for the dataset). Changes in inefficiency over time can also be captured—for example, in the half-normal case, if DNOs are moving away from the frontier, the variance would be increasing over time such that the mean inefficiency will also be increasing. Furthermore, if it is not possible to distinguish noise from inefficiency then the SFA model reduces to the standard OLS model with normal errors. This implies that the errors represent noise only and there is insufficient information to establish levels of relative inefficiency. (See section 4.1 for the results of the relevant statistical test.)

By making this assumption, SFA is able to decompose the residual term into inefficiency and noise, and thereby identify the relative inefficiency of each firm in the sample. SFA can also test for the presence of inefficiency.[18] The assumption on the distribution of inefficiency can be relatively flexible and its validity tested for. If the factors affecting inefficiency are observed, it is also possible to include them in the model as determinants of inefficiency, and, as such, the distributional assumptions on the inefficiency term can be made to be very flexible and their appropriateness can be tested for.[19] Thus, the application of SFA allows for the noise element and true inefficiency components, and would therefore reduce the amount of judgement required under Ofgem's current approach.[20] That is, while the distributional assumption for inefficiency represents an additional assumption for the estimation of inefficiency, the assumption can be tested for, alternative distributions can be used and the distributions can be made to be very flexible. This is far preferable to all the ad hoc adjustments that need to be made when using COLS, as well as the inconsistency issues.

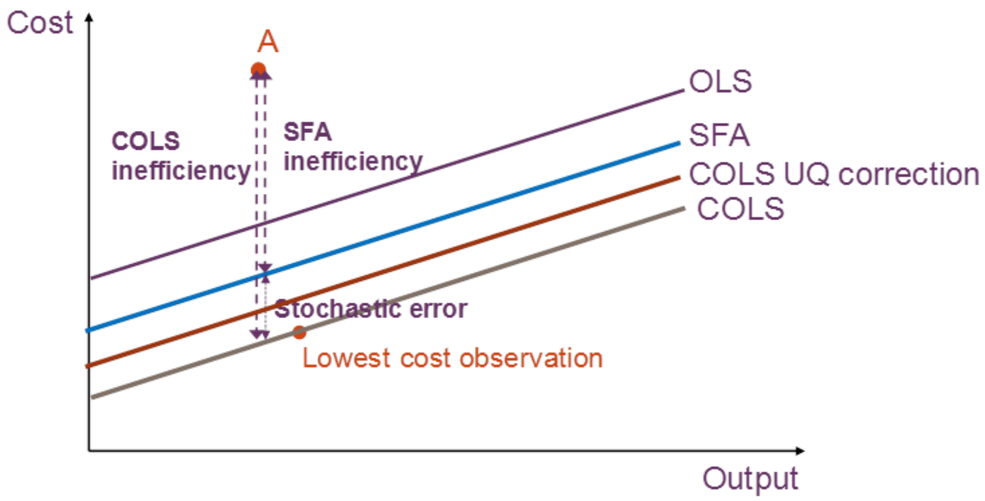
Figure 2.5 illustrates the difference between the COLS and SFA techniques for estimating inefficiency.

[18] See Kumbhakar and Knox Lovell (2000), op. cit.

[19] Ibid.

[20] Some degree of judgement is still necessary since an assumption needs to be made about the distribution of inefficiency, which, in practice, is unknown.

Figure 2.5 Estimating inefficiency using COLS and SFA



Note: The figure is intended for illustration purposes only; in particular, SFA explicitly accounts for noise in the data and does not assume that the same level of noise is present in all the observations, as in an upper-quartile or upper-third correction to the COLS inefficiency estimates.
Source: Oxera analysis.

The availability of panel data can greatly facilitate the application of SFA in the regulatory setting, in terms of both whether the approach can feasibly be applied in a specific setting—since having data over time increases the number of observations available—and the overall robustness of the approach. For example, if panel data is available, SFA can be undertaken without needing to assume beforehand the distribution of inefficiency (although this comes at the cost of the model producing a single inefficiency score for each comparator, regardless of the time period concerned—ie, the estimated inefficiency would be time-invariant).

Other SFA extensions and refinements made possible through the use of panel data include the 'true' random and fixed-effects models first suggested by Kumbhakar and Hjalmarsson in 1993 and further developed by Greene in 2005.[21] Such models can potentially account for both (time-invariant) company-specific factors and measurement error. As such, they can potentially control for discrepancies in the efficiency estimates across models that could arise owing to the specification of unobserved factors, such as topography, quality of service or the overall condition or quality of the capital base.

Recommendation 6: to reduce the subjectivity involved in using an upper quartile benchmark, Oxera would recommend the use of SFA, which takes into account symmetric errors/noise in efficiency modelling. With the current dataset, the panel data and SFA models considered in this report were found to be estimable (ie, practical). In addition, being a stochastic approach by explicitly accounting for data errors/noise in efficiency modelling, SFA enables statistical testing of the estimated efficiency levels in terms of calculating confidence intervals (ie, a confidence interval around the point estimate of inefficiency can be provided).

2.4.3 Estimating relative efficiency: interpretation of residuals

A key question is how to interpret the residuals from the modelling.

– Under OLS, the residuals explicitly represent noise. This is the foundation of OLS modelling and these residuals are assumed to have zero mean and constant variance

[21] See Kumbhakar, S.C. and Hjalmarsson, L. (1993), 'Technical efficiency and technical progress in Swedish dairy farms', chapter 9, pp. 257–70 in H.O. Fried, C.A. Knox Lovell and S.S. Schmidt (eds), *The Measurement of Productive Efficiency Techniques and Applications*, Oxford University Press; and Greene, W.H. (2005), 'Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model', *Journal of Econometrics*, **126**:2, pp. 269–303.

(symmetric and normally distributed for hypothesis testing). It is not possible to separate out inefficiency from this noise without some ad hoc assumption.

– As discussed above, panel data allows company-specific factors that cannot be observed or suitably measured, or variables that change over time but not across entities, to be controlled for. These company-specific factors can be assumed to represent either solely inefficiency or solely unobserved uncontrollable factors.

– By extending panel data analysis to include SFA, it is possible to avoid the two extreme assumptions that company-specific effects are either solely inefficiency or solely uncontrollable factors, and to separate noise from firm effects and thus identify persistent inefficiency and transient inefficiency (see section 3).

In these settings, the residuals have clear interpretations.

Recommendation 7: SFA allows for direct interpretation of the residuals. Indeed, further extensions in a panel setting allow for explicit interpretation in terms of company-specific effects, noise, persistent inefficiency and transient inefficiency. Similarly, panel data approaches (both fixed- and random-effects) are used in an efficiency context to allow direct interpretation of the residuals as inefficiency.

3 Panel models examined

Having examined cost benchmarking and some specific issues with regard to the dataset available and approaches proposed for RIIO-ED1, in this section, alternative efficiency modelling approaches are examined. These approaches are used in section 4 to estimate the relative inefficiency of the DNOs using the current dataset. However, before estimating the models, their advantages and disadvantages are discussed below.

3.1 Company-specific effects, noise, persistent inefficiency and transient inefficiency: overview

This section introduces a suite of models that can be used to estimate efficiency within a panel data context.

As explained in section 2.3, panel data allows for company-specific factors to be controlled for that cannot be observed or suitably measured (eg, differences in topography); or variables that change over time but not across entities (eg, regulatory changes). That is, panel data can account for individual heterogeneity. As such, panel data potentially allows differences to be captured between firms that are not explicitly modelled (known as firm heterogeneity or 'firm effects'), and that are not necessarily inefficiency.

As explained in section 2.4.2, the idea behind SFA is that it includes both a symmetric error term and an asymmetric inefficiency term. In order to separate these two elements, certain distributional assumptions are required. The distributional assumption on the inefficiency term is perhaps the element that receives the most criticism, but different distributional assumptions can be used (by including additional explanatory factors), and tested for.

Box 3.1 sets out the principles behind company-specific effects, noise, persistent inefficiency and transient inefficiency.

Box 3.1 Principles of the competing models

In general, the difference between the actual cost of a company and its estimated efficient cost from an econometric model can be broken down into **three components**: a **company-specific** or **firm effect**, **inefficiency**, and **noise**.

The company-specific effect (or firm effect) refers to factors that might give some companies a cost advantage (disadvantage) over their competitors. These factors are assumed to be different for different companies (ie, they represent the heterogeneity or differences between firms that may or may not be explicitly modelled for), but are likely to be invariant over time. While some of them (eg, population density or regional wages) can be explicitly accounted for in the model, others (eg, topology, condition of network, or quality of service) are often *unobserved* (or cannot be suitably quantified), and are thus unaccounted for in the model. If these effects are company-specific but invariant over time, their joint effects are captured by the intercept term in the regression, which is referred to as 'firm heterogeneity' or 'firm effects'.

When data on companies over time is available, within the modelling framework, the 'panel' structure of the dataset can either be ignored (thereby 'pooling' the data, as preferred by Ofgem), or it can be explicitly recognised that the data represents information on the same companies over time.

Where the data is pooled together and the model estimated using COLS regression, all three components are treated as inefficiency. (Ofgem attempts to adjust for noise by using an upper-quartile or upper-third benchmark, although, as discussed above, this approach is somewhat ad hoc and assumes that the 'noise' is equal for each company.) In a panel framework, it is possible to separate inefficiency from the other two components.

Having data over time on the companies can help to control for the *unobserved* company-specific factors. For example, a common approach in a panel set-up to control for these factors is to use company-specific dummies in the OLS regression (this is the company fixed-effects approach). However, this approach may not be appropriate when some of the explanatory variables do not change much over time (eg, population density).

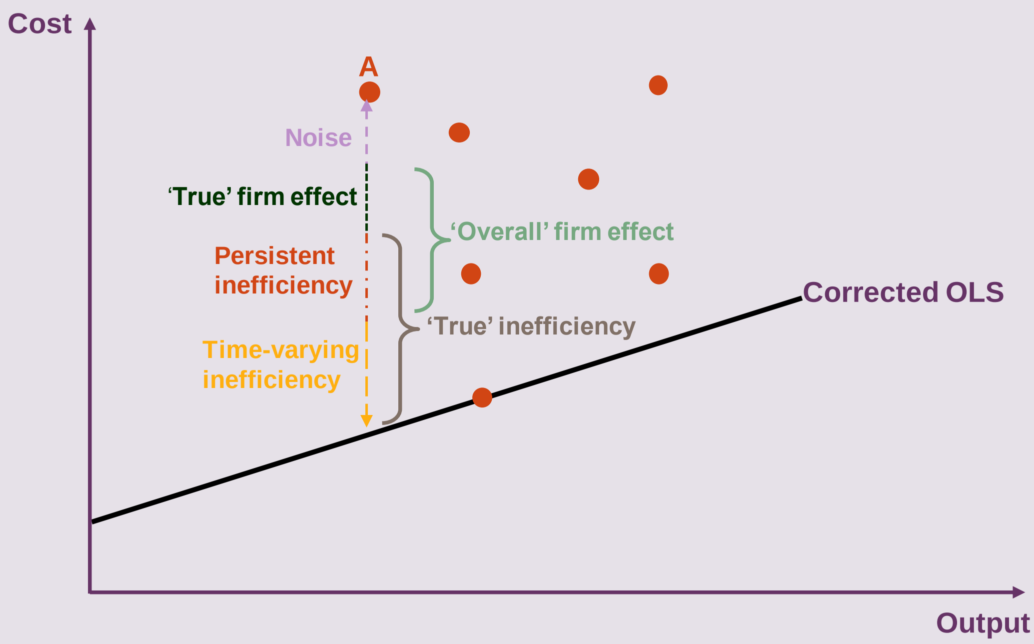
In the context of efficiency analysis, the models developed in a panel framework controlling for company-specific effects differ in terms of the assumption made about these effects—ie, whether they can be considered to represent solely inefficiency or solely unobserved uncontrollable factors (and thus not be part of inefficiency). However, neither 'extreme' assumption can be considered entirely satisfactory, as the overall effect could be a mix of both—it is more likely that there is a component that captures the 'true' firm effects and a residual term that captures inefficiency. The SFA approach can help to separate the inefficiency component from the 'true' firm effects.

Separate from the firm-specific effects that may or may not be inefficiency, companies' efficiency levels can change over time. This component is commonly referred to as transient, or time-varying, inefficiency. The SFA approach can once again help to capture and identify this component.

The overall efficiency of the company is then a product of the *time-invariant* efficiency term obtained from the firm effect and a *transient* efficiency term, with both components estimated using SFA.

The breakdown of the three components is illustrated in the figure below.

Graphical illustration of the breakdown of the total gap



Note: This figure is intended only as an illustration of the breakdown of the overall gap between actual cost and the estimated efficient cost into three components: a firm effect, noise, and inefficiency. True firm effects and noise are both positive in the figure.
Source: Oxera.

The models considered in this section include those based on SFA as well as standard panel data models, both of which improve on pooled OLS regression as they aim to separate noise and/or company-specific effects from inefficiency.

3.2 Models that do not take into account that there are observations over time

3.2.1 Baseline model: Ofgem's 'preferred' pooled OLS

This model simply pools the data and assumes that each observation represents an independent company, as it ignores the fact that the data consists of repeated observations on the same companies over time. No explicit account is taken for noise in the model, although, in the past, Ofgem has applied an ad hoc upper-quartile (or upper-third) benchmark when using these models. An upper-quartile benchmark would reduce the efficiency gap for all companies to the frontier by defining the frontier as being between the third and fourth most efficient companies, rather than the first.

As discussed in section 2.3, ignoring firm effects might give inconsistent estimates of the parameters of the cost drivers, and incorrect estimates of inefficiency that are based on the residuals (regardless of how inefficiency is measured).

3.3 Panel data—models with time-invariant inefficiency

3.3.1 Model 1: each company's inefficiency is assumed to be the same over time (time-invariant)

Model 1(a): Without distributional assumptions

Panel data plays an important role in efficiency estimation. In the early panel models, inefficiency was assumed to be firm-specific and time-invariant. In these cases, the unobservable firm effects are treated as either random variables or fixed parameters. Efficiency is defined relative to the best firm, which is taken to be as 100% efficient.

A major problem with the fixed-effects model is that it cannot separate inefficiency from time-invariant explanatory variables—the effects of all fixed observed and unobserved explanatory variables will be captured by the firm-specific intercepts. (This is important in the current context, as many explanatory variables are time-invariant or close to being time-invariant, and, indeed, as a result, the fixed-effects approach does not work with the current dataset.) This can be corrected in a random-effects model in which the firm effects are viewed as a random variable, independent of cost drivers and noise. In a similar way to the fixed-effects model, efficiency can be estimated relative to the best firm in the sample.

Model 1(b): With distributional assumptions

Alternatively, SFA can be used, which requires distributional assumptions for both noise and inefficiency. An advantage of this approach is that time-invariant explanatory variables (here, cost drivers) can be used in the cost function. Furthermore, efficiency estimates are not relative to the best firm in the sample (and the best firm can be inefficient as well).

Note that these models assume that inefficiency is time-invariant, which is unrealistic in most situations, although it may be appropriate if only a short time period is being modelled.

3.4 Panel data—models that allow efficiency to change over time

3.4.1 Model 2: time-varying inefficiency (without distributional assumptions)

The next generation of models relaxed the assumption that inefficiency is time-invariant. This is achieved with and without distributional assumptions.

In the model without distributional assumptions, the fixed- or random-effects model is extended by allowing these effects for each firm to change over time. For example, it can be modelled as a quadratic function of time for each firm, thereby allowing the temporal pattern of inefficiency to be different for different firms. Inefficiency can then be defined relative to the best firm, and these inefficiencies will change over time in a flexible manner. This approach has the disadvantage that it has too many firm-specific parameters to be practically estimated in a dataset for which the number of time periods is small. A restrictive version of it would be to allow the temporal pattern of inefficiency to be the same for all firms.

3.4.2 Model 3: time-varying inefficiency in SFA models (ie, with distributional assumptions)

In the model with distributional assumptions (ie, SFA), Model 1 can be similarly extended to allow inefficiency to change over time by allowing the mean and/or variance functions of the inefficiency component to change over time. A further generalisation has both the mean and the variance of inefficiency as functions of exogenous variables (eg, company-specific factors), which can be time-varying. (Such a model can be insightful for operational purposes by exploring the impact of (controllable) drivers on inefficiency.)

Time-varying inefficiency can also be explicitly introduced in the SFA framework in many other ways. In one popular model, inefficiency, u_{it}, is specified as the product of a time-invariant random (stochastic) component, u_i, and a time-varying deterministic component, g(t)—ie,

$$u_{it} = g(t) \, . \, u_i$$

where u_i is a one-sided random variable, and t is time. Time dependence is introduced via the g(t) function, and both the mean and the variance of inefficiency are functions of g(t), which is often labelled the ‘scaling function’. The g(t) function determines the temporal behaviour of inefficiency and is assumed to be the same for all firms. This model has the advantage that it can:

– accommodate time-invariant explanatory variables;

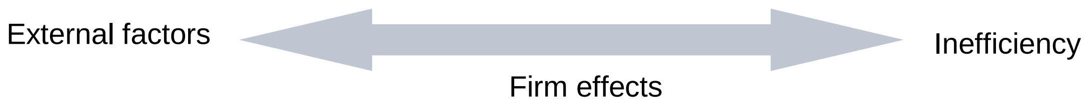
– separate time-varying inefficiency (ie, catch-up) from technical change (ie, shifts in the frontier).

In this model, time plays two roles. It allows the frontier to shift by including time as a cost driver, and the inefficient firms to move towards the frontier (or away from it) via the g(t) function (which dictates the temporal pattern). This model can be further generalised to allow the temporal pattern of inefficiency to be firm-specific, by allowing g(t) to be $g(z_{it})$—ie, $u_{it} = g(z_{it})\ u_i$ (where subscript i denotes the firm and t denotes time).

It is worth noting that these SFA models can accommodate firm effects that can be treated as fixed or random (as in the fixed-effects and random-effects panel models).

A fundamental and perhaps philosophical question is whether the firm effects (fixed or random) are capturing parts of inefficiency that is persistent (time-invariant). None of these models explicitly separates firm effects from time-invariant inefficiency. The problem comes from the interpretation of firm heterogeneity—ie, whether it constitutes time-invariant external factors or inefficiency. The two extreme assumptions are shown in Figure 3.1.

Figure 3.1 Two extreme assumptions for the interpretation of firm effects



Source: Oxera.

Model 1 did not distinguish between these two assumptions because firm effects were viewed as unobserved management effects, and therefore regarded as inefficiency. However, it is not clear whether all, or just parts, of the firm effects are inefficiency. A related issue is whether there is an inefficiency component that is time-varying which is not captured in the models.

Given this backdrop, the next generation of models followed two opposite paths (Model 4 and Model 5).

3.5 Models that take into account company-specific effects

3.5.1 Model 4: time-varying inefficiency, where firm effects are considered to represent external factors

This group of models views firm effects (fixed or random) as something different from inefficiency. That is, inefficiency in these models is always time-varying (and can be a function of exogenous variables). Thus, the models assume that firm effects relate to exogenous characteristics not captured in the model, and that they are not related to inefficiency (ie, the left-hand side of Figure 3.1). As such, these models ignore the possibility of persistent inefficiency, which is instead hidden within the firm effects.

3.5.2 Model 5: time-varying inefficiency models with persistent inefficiency

The alternative group of models treat firm effects as persistent inefficiency (ie, the right-hand side of Figure 3.1), and add a second component to capture time-varying inefficiency. That is, in these models firm effects (that are not part of inefficiency) are ignored and are confounded in persistent inefficiency.

Consequently, Model 4 and Model 5 are both mis-specified, although their impacts on estimated inefficiency are not the same. Model 4 is likely to **under-estimate** overall inefficiency, especially if persistent inefficiency exists. In contrast, Model 5 is likely to **overestimate** inefficiency by treating firm effects as solely due to inefficiency. The actual inefficiency is likely to be somewhere between these two extremes. As such, the final modelling framework seeks to address this issue.

3.6 Models that take into account company-specific effects, time-varying inefficiency, firm effects, and persistent and transient inefficiency

3.6.1 Model 6: time-varying inefficiency, with firm effects, persistent and transient inefficiency

Model 6 overcomes some of the limitations of the earlier models. The error term is split into four components to take into account different factors affecting cost, given the cost drivers:

– the first component captures the firm's heterogeneity (firm effects), which has to be disentangled from persistent inefficiency effects (which are assumed to be time-invariant);
– the second component captures persistent (or time-invariant) inefficiency;[22]
– the third component captures time-varying inefficiency;[23]
– the last component captures noise or random shocks.

3.7 Summary

The components of the gap between actual cost and efficient cost that are viewed as inefficiency under the different models are illustrated in Table 3.1 below. The 'true' model is assumed to have both time-invariant (persistent) and time-varying (transient) inefficiency, in addition to company-specific effects and noise terms.

[22] For example, inefficiency is associated with (unobserved) management, and management is assumed to be time-invariant.

[23] It is probably more realistic to assume that management changes over time, although most of it might be time-invariant. That is, management has a time-invariant and a time-varying component. If, as before, inefficiency is associated with management, there is a situation in which one part of inefficiency is time-invariant and the other part is time-varying.

Table 3.1 Taxonomy of the efficiency models on the basis of what is viewed as inefficiency

Identified as	Components of the gap between actual and efficient cost: True firm effect	Time-varying inefficiency	Time-invariant inefficiency	Noise
Pooled COLS	Inefficiency	Inefficiency	Inefficiency	Inefficiency
Model 1(a): random-effects	Inefficiency		Inefficiency	
Model 1(b): time-invariant SFA	Inefficiency		Inefficiency	
Model 2: time-varying SFA		Inefficiency		
Model 3: SFA with scaling function		Inefficiency		
Model 4: firm-effects SFA		Inefficiency		
Model 5: firm-effects with time-varying SFA	Inefficiency	Inefficiency	Inefficiency	
Model 6: four-component model		Inefficiency	Inefficiency	

Note: The table is intended only as an illustration; the components are not necessarily separable under the different models.
Source: Oxera.

In pooled COLS all the components are treated as inefficiency, in the sense that the residual which contains all the four components is compared against the minimum residual to estimate inefficiency. In Models 1a and 1b, inefficiency is viewed as time-invariant and therefore captures both firm effects and time-invariant effects (these two cannot be separated). In Models 2 and 3, inefficiency is viewed as time-varying and everything else is dumped into the noise term. In Model 4, firm effects are treated as not part of inefficiency and only time-varying inefficiency is captured. Thus, in Models 2–4 only time-varying inefficiency is captured. Since these models differ in other respects, the resulting inefficiency estimates are not the same. Some of the models could be mis-specified, which might result in biased parameter estimates and incorrect residuals. Model 5 treats firm effects as part of inefficiency and identifies firm effects, time-varying and time-invariant inefficiency. Model 6 separates time-invariant inefficiency from firm effects, and captures both persistent and time-varying inefficiency.

The main characteristics of the panel data models are summarised in Table 3.2.

Table 3.2 Main characteristics of the panel data models

	Pooled COLS	Model 1: (a) random-effects (b) time-invariant SFA	Model 2: time-varying SFA	Model 3: SFA with scaling function	Model 4: firm-effects SFA	Model 5: firm effects with time-varying SFA	Model 6: four-component model
Firm effect	No	Yes (treated as inefficiency)	No	Yes (fixed)	Yes (random)	No	Yes (random)
Technical inefficiency							
Persistent	No	No	No	No	No	Yes	Yes
Transient	No	No	Yes	No	Yes	Yes	Yes

Source: Oxera.

The pros and cons of the competing models are summarised in Table 3.3 below.

Table 3.3 Pros and cons of competing models

Models	Pros	Cons	Comments
Pooled OLS	Simplest of the models	Ignores the fact that the data consists of repeated observations on the same companies over time Assumes that the gap between actual cost and that predicted by the OLS model is inefficiency, and thus confounds inefficiency with noise Since noise is ignored, inefficiency being defined relative to the best firm (defined over all firms and time) is overestimated.	Regulators take account of company factors outside the modelling by normalising the cost for various company-specific factors. These involve regulatory judgement and uncertainty, and may reduce the accuracy of the model Regulators often apply some form of ad hoc adjustment to the efficiency scores estimates from COLS to account for noise, although this correction assumes that noise is uniform across observations (this assumption contradicts the fact that noise is random) In DPCR5, while Ofgem indicated that it would use COLS to determine the efficiency scores (which require shifting the intercept to the upper-quartile or upper-third level), it appears that it adjusted both the slope and the intercept estimates, which could be considered as non-standard and atheoretical
Model 1, (a) random-effects, (b) time-invariant SFA	Accounts for the data being multiple observations of the same firm over time Separates inefficiency from noise	Assumes that inefficiency is time-invariant, which may not be true and, if the time period of the panel dataset is relatively long, this may not provide a good indication of a company's current inefficiency Cannot accommodate determinants of inefficiency that are time-varying) Cannot distinguish between persistent inefficiency and time-invariant cost drivers (fixed-effects)	Ofwat is currently considering random-effects and SFA (both time-invariant and time-varying) as possible options for PR14 Ofcom has considered both random-effects and SFA to assess BT's efficiency The ORR used SFA to assess the efficiency of Network Rail in PR08 and for the current review
Model 2, time-varying SFA	Allows inefficiency to change over time Accounts for the data being multiple observations of the same firm over time Model can separate time-varying inefficiency (catch-up) from technical change (shift in the frontier) Separates inefficiency from noise	Assumes that inefficiency changes over time via the mean and/or variance of inefficiency, which may be related to exogenous factors Cannot distinguish firm effects from persistent efficiency	Ofwat is currently considering random-effects and SFA (both time-invariant and time-varying) as possible options Ofcom has considered both random-effects and SFA to assess BT's efficiency The ORR used SFA to assess the efficiency of Network Rail in PR08 and for the current review

Models	Pros	Cons	Comments
Model 3, SFA with scaling function	Model can accommodate time-invariant explanatory variables Model can separate time-varying inefficiency (catch-up) from technical change (shift in the frontier) Accounts for the data being multiple observations of the same firm over time Separates inefficiency from noise	Cannot distinguish external factors from persistent efficiency	
Model 4, firm-effects SFA	Model can identify time-varying inefficiency Accounts for firm-specific external factors Accounts for the data being multiple observations of the same firm over time Separates inefficiency from noise	Assumes there is no persistent inefficiency component, which there might be Underestimates overall inefficiency by ignoring persistent inefficiency	Numerous empirical studies of the model in different sectors and jurisdictions have been published in the literature Could be used to provide a lower bound of the estimated cost-reduction targets or could be directly applied, since, by being a lower-bound estimate, it could provide an incentive for companies to outperform
Model 5, random-effects with time-varying SFA	Model can identify time-varying inefficiency Assumes there is a persistent inefficiency component Accounts for the data being multiple observations of the same firm over time Separates inefficiency from noise	Cannot distinguish external factors from persistent efficiency Overestimates overall inefficiency by treating firm effects as inefficiency	Could be used to provide an upper bound of the estimated cost-reduction targets
Model 6, four-component model	Model can identify time-varying inefficiency Assumes there is a persistent inefficiency component Can distinguish between external factors and persistent efficiency Can be relatively robust to the inclusion/exclusion of company-specific factors in the function Accounts for the data being multiple observations of the same firm over time Separates inefficiency from noise	Although the modelling approach is relatively more robust and flexible, it could be argued to be more involved compared with pooled OLS	

Source: Oxera.

4 Practical application: estimation results

Oxera's analysis is based on the dataset used by Frontier Economics, which was provided by ENWL. In each model, the specification of the cost function used by Frontier Economics is considered—ie, TOTEX is used as the cost measure, and the cost drivers are number of consumers, peak capacity, population density, a real wage index and a time trend.

The discussion in this section is limited to the results from the different panel modelling techniques discussed in section 5. In its analysis, Frontier Economics has indicated data issues with the cost data for the first two years and some of the cost drivers. These issues were not considered in this analysis.

4.1 Statistical robustness of the panel models

Empirical results indicate that the panel models considered in the report have reasonable statistical properties and enable sound economic interpretation of the regression coefficients.

Table 4.1 summarises the results using two statistical tests which test the assumptions behind the panel models (against the pooled OLS model) on the current dataset (in addition, all the models pass the Wald test[24]). These are the likelihood ratio (LR) test and the Lagrange multiplier (LM) test. If the null hypothesis is rejected in these tests (here, fixed at 5%), the modelling assumptions are statistically sound and the models are preferred to OLS. The table also shows whether the estimated coefficients are intuitive.

For the SFA models, the LR tests for the absence of technical inefficiency in the model. If this is found to be the case, the SFA model reduces to the standard OLS model with normal errors. In Models 4 and 5, the LR test is applied to the composed error term estimated using a random effects model; and in Model 6, two LR tests are applied—one on the company-specific component and another on the composed error term, with both components estimated using a random effects model. For the random-effects model, a specific LM test is applied where, if the null hypothesis of no random effects is rejected, an OLS model is deemed not to be appropriate.

Once the data and specification issues identified by Frontier Economics are addressed, these tests will warrant additional investigation.

[24] This tests for the null hypothesis that the coefficients of the explanatory variables (here, number of consumers, peak capacity, population density, population real wage index and a time trend) are simultaneously equal to zero.

Table 4.1 Summary of the results of statistical tests

	Does the model pass the test at 5% significance level? Likelihood ratio test	Lagrange multiplier test	Does the model enable sound interpretation of the coefficients?
Model 1(a): random-effects	n/a	Yes	Yes
Model 1(b): time-invariant SFA	Yes	n/a	Yes
Model 2: time-varying SFA	Yes	n/a	Yes
Model 3: SFA with scaling function	Yes	n/a	No
Model 4: firm-effects SFA	No	Yes	Yes
Model 5: random-effects with time-varying SFA	No	Yes	Yes
Model 6: four-component model	No, Yes	Yes	Yes

Note: For each test, if the probability value is below 0.05 (or 5%), the 'null hypothesis' of the test is rejected. Here, if the null hypothesis is rejected, the assumptions of the model are statistically sound; in all the models, the estimated coefficients are jointly significant at 5%, thereby passing the Wald tests (not shown). Results from the fixed-effects model are not presented here as it is inappropriate for the current application, where some of the explanatory variables are time-invariant. That random effects are appropriate is also confirmed by the Hausman test, where the null hypothesis is random effects and the alternative hypothesis is fixed effects (the p-value of the test is 0.55, which indicates that the null hypothesis of random effects cannot be rejected).
Source: Oxera.

Models 4, 5 and 6 control for firm effects using the random effects model, which the LM test confirms is appropriate. As discussed in sections 3.5 and 3.6, these three models make different assumptions on the company-specific effect and the noise component estimated by the random effects model. In Model 4, company-specific effects are included but are not part of inefficiency, and the presence of inefficiency is tested in the time-varying composed error term (noise plus inefficiency) using SFA.

In contrast, Model 5 assumes that firm effects are part of inefficiency but similar to Model 4, tests for the presence for inefficiency in the time-varying composed error term (noise plus inefficiency) using SFA. Finally, Model 6 makes no such extreme assumptions and subjects both the firm effect and the noise component to SFA to separate time-invariant inefficiency from the 'true' firm effect and transient inefficiency from the composed noise component. On the current dataset, SFA identifies firm effects as inefficiency, but that no transient inefficiency is present in the composed error term, and hence one of the LR tests (ie, for transient inefficiency) is rejected. However, when the data and specification issues are addressed or when forecast data is being used, SFA could identify inefficiency in both components, and this would need to be tested.

The test results on the interim data indicate that the assumptions of all the panel models are robust (in Model 3, the estimated coefficients are not intuitive, which might warrant additional examination when the data and specification issues are addressed).

The key and consistent outcome of the statistical tests is that pooled OLS is inappropriate for the dataset, and Ofgem should consider the panel models discussed in this report as part of its cost assessment toolkit for RIIO-ED1, as they appear to be more appropriate for the current dataset.

4.2 Results from the panel models

The companies' efficiency scores at the industry level under each model are summarised in Table 4.2, and Figure 4.1 shows the average efficiency gap to the frontier graphically.

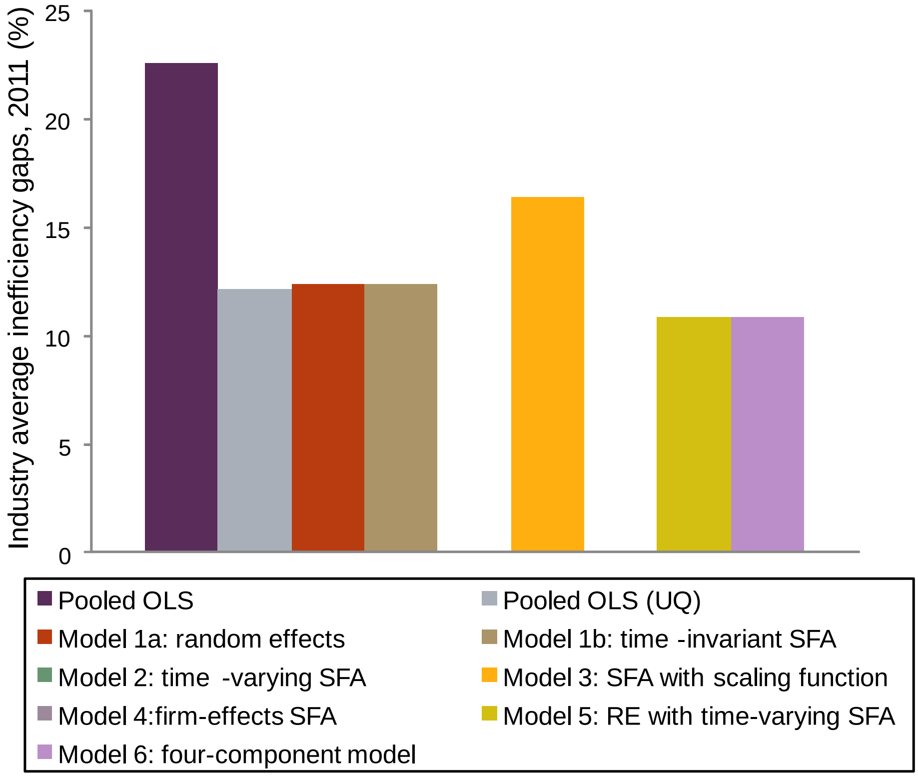
For Model 1(a) and 1(b), inefficiency is assumed not to change over time, such that the estimated inefficiency for each DNO represents, in some sense, its average performance over the period. The other models provide an inefficiency estimate for each year in the panel dataset (although only the 2011 figure is provided in the table and figure).

Table 4.2 Summary of the efficiency position under different models (%)

	Average inefficiency	Minimum DNO inefficiency	Maximum DNO inefficiency	Standard deviation of the efficiency gaps
Pooled OLS	23	0	40	12
Pooled OLS (upper-quartile benchmark)	12	0	30	10
Model 1(a): random-effects	12	0	24	8
Model 1(b): time-invariant SFA	12	3	22	7
Model 2: time-varying SFA	0	0	0	0
Model 3: SFA with scaling function	16	4	36	12
Model 4: firm-effects SFA	0	0	0	0
Model 5: random-effects with time-varying SFA	11	0	21	7
Model 6: four-component model	11	0	21	7

Note: Where appropriate, all figures are relative to the frontier/benchmark in the year 2011. The efficiency scores are rounded to the nearest integer value. The results from the company fixed-effects model are not presented here, as the model is unable to distinguish between time-invariant cost drivers and time-invariant inefficiency.
Source: Oxera analysis.

Figure 4.1 Illustration of the efficiency position under different models



Note: Where appropriate, all figures are relative to the frontier/benchmark in the year 2011. The efficiency scores are rounded to the nearest integer value. The results from the company fixed-effects model are not presented here, as the model is unable to distinguish between time-invariant cost drivers and time-invariant inefficiency.
Source: Oxera analysis.

Consistent with the theoretical arguments presented in section 5, pooled OLS tends to overestimate inefficiency, as it cannot separate inefficiency from company-specific effects and noise in the data or modelling errors. In this instance, this overestimation is significant, as seen in the table of results, where the average gap is estimated to be 23%—around double the estimates from the majority of other approaches. As such, pooled OLS is inappropriate.

However, with the use of a subjective upper-quartile benchmark, the industry average inefficiency reduces to 12%, which, *in this dataset*, is similar to that estimated from some of the other models, including the random-effects model. As such, the ad hoc adjustment is, on average, approximately consistent with the other approaches in this case. This could be coincidence and would need to be checked in other circumstances (ie, when the data is revised and, more importantly, on the forecast data). However, while at the industry level the average gap is similar under the two models, at the DNO level there is considerable variation in the estimated gaps. This can be seen, for example, from the spread of the estimated efficiency scores (as measured by the range and standard deviation), which under the random-effects model and most of the SFA models is relatively more compact than that under the upper quartile-adjusted COLS approach.

The ad hoc nature of the upper-quartile adjustment is further illustrated in Table 4.3. This shows that, while the adjustment could be reasonable for some companies, even at the DNO level (which again could be a coincidence with this dataset), in the majority of instances with this dataset, the adjustment significantly under- or over-corrects the estimated gap for noise.

Table 4.3 Inefficiency gap of some DNOs in 2011 by model (%)

	Reasonable		Over-correction		Under-correction		Reasonable
	NPGY	WMID	LPN	SSEH	SPN	SWales	Industry average
Pooled OLS	7	30	21	22	40	25	23
Pooled OLS (upper-quartile benchmark)	0	19	8	9	30	13	12
Model 1: random-effects	0	20	14	19	19	5	12
Model 1: time-invariant SFA	3	19	13	18	18	6	12
Model 2: time-varying SFA	0	0	0	0	0	0	0
Model 3: SFA with g(t)	4	28	12	14	35	7	16
Model 4: firm-effects SFA	0	0	0	0	0	0	0
Model 5: random-effects with time-varying SFA	0	18	12	16	17	4	11
Model 6: four-component model	0	18	12	16	17	4	11

Note: Where appropriate, all figures are relative to the frontier/benchmark in the year 2011. The efficiency scores are rounded to the nearest integer value.
Source: Oxera analysis.

As seen from Table 4.3, while NPGY's and WMID's estimated gaps to the upper-quartile benchmark are comparable to those estimated from the panel data and SFA models, the adjustment over-corrects for noise by about 5–10% in the case of LPN and SSEH, and under-corrects by about 8–11% in the case of SPN and S Wales. At the TOTEX level, such discrepancies could be significant in monetary values.

Table 4.3 demonstrates the inconsistency and risks in applying a uniform upper-quartile adjustment to the estimated gap across all DNOs, thereby not recognising the differing levels of noise across the companies. Together with the limitations of the pooled OLS when company-specific effects are not accurately controlled for (which, as Ofwat and its academic advisers argue, are often difficult to quantify accurately using a regulator–company dialogue), the accuracy of the estimated inefficiency gaps is questionable and could constitute a significant regulatory risk.

Among the panel models, the random-effects estimate of inefficiency is very similar to that of the time-invariant SFA models. This is expected, as they take a similar approach to distinguishing between inefficiency and noise. The range of the estimated efficiency scores is also compact under random-effects and time-invariant SFA when compared with COLS.

As discussed in section 4.1, the coefficients from Model 3 are not intuitive on the current dataset and might warrant additional examination once the data and specification errors are addressed.

Among the models that separate company-specific effects from inefficiency, Model 4 (firm-effects SFA) is expected to form a lower bound of inefficiency, tending to overestimate efficiency by 'hiding' persistent inefficiency within firm effects and measuring only time-varying inefficiency. Inefficiency in Model 4 is just the time-varying portion, and the model (similar to time-varying SFA) is unable to identify a statistically significant amount of inefficiency (ie, all units are determined to be 100% efficient).

Likewise, according to the theory discussed in section 3, Model 5 (random-effects with time-varying SFA) tends to overestimate inefficiency as it attributes firm effects to inefficiency. The model estimates average inefficiency to be 11% (similar to random effects), but calculating a smaller gap to the frontier than pooled OLS.

Finally, Model 6 demonstrates that, when time-varying inefficiency, persistent inefficiency, firm effects and noise are able to be separately estimated, the average inefficiency gap is estimated to be 11%, which is higher than that estimated using the random-effects model or the time-invariant SFA model.[25]

On the current dataset and for the specification employed, the results from Models 5 and 6 are the same, as there is no residual inefficiency and the only inefficiency that is estimated to be present is persistent inefficiency (consistent with the results of the statistical tests discussed in section 4.1). In addition, the persistent inefficiency estimated from the models is similar, as the firm effects can be attributed to inefficiency. However, this is likely to change on the forecast data, where, for example, companies could make different assumptions on when and where (ie, which cost area) to take out costs, and the results from the two models could be different (as there is likely to be some residual inefficiency present in the companies).

Finally, consistent with the theoretical discussion in section 3, the overall patterns of the efficiency scores under the different models are largely similar across the companies; in particular, Model 6 improves on the efficiency scores estimated for the companies relative to pooled OLS and random effects.

It is worth emphasising that all the models estimated in this report can be implemented using publicly available software tools—indeed, the most flexible model, Model 6, uses a multi-step approach and can be implemented with the current version of Stata, without the need for additional code. Moreover, they are all established in the academic literature.[26]

[25] The efficiency position of all the companies under different modelling approaches can be seen in the spreadsheet accompanying this report.

[26] The modelling for this project was undertaken in Stata and the codes for the different models are used with the authors' permission, and come from chapter 10 of Kumbhakar, S., Wang, H.-J. and Horncastle, A. (forthcoming), *Practitioner's Guide to Stochastic Frontier Analysis*, Cambridge University Press.

5 Conclusion

Ofgem has stated its intention to use a toolkit approach to determine the DNOs' relative efficiency for RIIO-ED1. Despite this, it is understood that Ofgem recently stated that, unless there is a rationale for opting for approaches that could control for company-specific factors within modelling, it is currently minded to focus on pooled OLS as the econometric method to estimate the DNOs' efficiency levels.

This report has demonstrated, both theoretically and empirically, that pooled OLS leads to incorrect conclusions when there are company-specific effects, as it overestimates inefficiency. With the current dataset, this overestimation is of the order of doubling the industry average inefficiency gap. As such, Oxera has recommended examining various panel data and SFA modelling approaches in comparison with the pooled OLS model. In all the models, the OLS model is a special case and therefore the OLS specification can be statistically tested.

To derive more reasonable inefficiency estimates, Ofgem applies ad hoc adjustments to the estimated inefficiency gaps from pooled OLS. It has previously used an upper-quartile or upper-third benchmark to determine the gap, depending on the range of the estimated gaps. However, the upper-quartile correction is ad hoc as it assumes that the same level of noise is present in the data across all the companies, thereby potentially under- or over-correcting for some DNOs. This is demonstrated empirically with the current dataset where, while at the industry level, the upper-quartile benchmark provides a reasonable correction, at the DNO level, in most instances, the inefficiency gap is either under- or overestimated. In particular, on the current dataset, eight DNOs' efficiency gaps are over-corrected for noise with a maximum difference of +7% and six DNOs' efficiency gaps are under-corrected for noise with a maximum difference of –13%, with two DNOs' differences in the gaps less than ±1%. At the TOTEX level, this subjective adjustment could result in considerable discrepancies in monetary terms, and, as such, could constitute a significant financial risk. In addition, Ofgem's current approach necessitates regulatory judgement to normalise for DNO-specific differences. As the sizes of the cost adjustments are likely to be difficult to quantify, this represents an additional regulatory risk and may further reduce the accuracy of the cost assessment approach.

In contrast, there is a wealth of literature on efficiency estimation within the panel framework that can reduce this need for ad hoc assumptions and unnecessary regulatory judgement.[27] These alternative models provide more robust estimates of the DNOs' relative efficiency levels and hence improve Ofgem's efficiency analysis. The approaches that provide the most accurate assessments could also be used to establish the extent to which a simpler approach requires company-specific adjustments in order to account for modelling errors and company-specific effects.

Based on theoretical foundations and empirical analysis (which confirms the theoretical insights), this report provides a number of findings and recommendations. These recommendations will result in a more robust estimation of the DNOs' relative efficiency levels than the use of pooled OLS alone, or could be used to provide a more robust basis for establishing ad hoc adjustments for each DNO.

– It is shown that OLS estimates can be biased and thereby result in incorrect estimates of inefficiency (no matter how inefficiency is calculated), because the residuals on which inefficiency measures are based are incorrect.

[27] For a survey of the different panel models, see chapters 3, 4, 6 and 7 of Kumbhakar and Knox Lovell (2000), op. cit.

– SFA and DEA have regulatory precedent in the UK and across Europe. To increase the robustness of the results, Oxera recommends that a number of approaches are examined, including SFA and DEA. In addition, all the models estimated in this report can be implemented using publicly available software tools and are well established in the academic literature.

– The results from the different approaches should be compared and contrasted, and, based on an understanding of the approaches some consensus could be reached in order to identify a more robust range for the estimated inefficiencies.

– SFA and panel data models allow for direct interpretation of the residuals. In all the SFA models, the statistical significance of inefficiency can be tested. This is not the case with OLS (or COLS). Indeed, further extensions in an SFA panel setting allow for explicit interpretation of the results in terms of uncontrollable company-specific effects, noise in data/modelling errors, persistent inefficiency and transient inefficiency. Such a decomposition and interpretation is currently not possible using other approaches.

– The alternative models examined in this report have been shown to be both valid and practical for the current dataset in terms of statistical robustness and economic interpretation of the results from the model.

– While Oxera would recommend examining panel data modelling approaches, it also appears that the fixed-effects panel model is not appropriate in this dataset, as some of the cost drivers do not change much over time. A random-effects approach would therefore be more appropriate. This is also empirically determined on the current dataset using a statistical test.

– In this instance, with the dataset examined in this report, the ad hoc adjustment of an upper-quartile is close, at an industry-wide level, to making an appropriate adjustment for errors. This could be coincidence and would need to be checked in other circumstances.

– However, while the average gap is similar to that of other models following this adjustment at the industry level, at an individual DNO level there is considerable variation in the estimated gaps—in some cases, the adjustment is too much and in others it is too little. At the TOTEX level, such differences could be significant in monetary terms.

Finally, in this report it has only been possible to examine historical costs and cost drivers empirically. Ofgem may, however, place weight on efficiency assessments made using companies' forecast data, as it did in RIIO-GD1. Such forecast data (over a period of eight years) is likely to show far more variation over time than actual historical data, and, in such a context, separating efficiency that does not change over time from transient inefficiency, and from errors and company-specific effects, may be more important than in the current context.

A1 Different approaches to cost benchmarking by UK regulators

This appendix summarises the approaches to cost benchmarking adopted by UK regulators (such as Ofgem, Ofwat, the ORR and Ofcom). According to recent surveys on international best practice on cost benchmarking in a regulatory context, SFA and DEA rank as the most commonly used approaches.[28]

A1.1 Ofgem's approach(es)

A1.1.1 DPCR5

In DPCR5, Ofgem collated data on the operational activities (OPEX) of each DNO over four years (2005/06–2008/09), pooled the data and used an OLS technique with year dummies to estimate the efficient expenditure levels. The year dummies were included to capture the time effect on the average costs of the activities in each year. Ofgem also considered a DNO-specific fixed-effects model and noted that the technique would enable ease of replication. However, Ofgem limited its assessment to using pooled OLS on the panel data for reasons of 'greater transparency' and 'data constraints'. Separate from the regression model, Ofgem undertook 27 DNO-specific cost adjustments (17 cost exclusions and ten normalisations) made on a case-by-case basis to take account of factors that are outside the control of the DNOs but have an impact on their cost performance. CAPEX was assessed using bottom-up engineering analysis.[29]

A1.1.2 GDPCR1

In GDPCR1, Ofgem used data in a single year (2006/07) to assess direct OPEX using an OLS technique. Similar to DPCR5, CAPEX was assessed using bottom-up engineering analysis.[30]

A1.1.3 RIIO-GD1

In RIIO-GD1, while the assessment approach was limited to pooled OLS on a panel dataset with year dummies, the analysis was extended to all the cost areas—ie, CAPEX, REPEX and TOTEX—apart from OPEX. While the previous reviews in the energy sector were based on assessment using historical data only, Ofgem put weight on efficiency assessments made using companies' forecast data as well. Although Ofgem had eight years of forecast data to model—forecast data showed far more variation over time as companies had assumed different cost profiles over the period in the business plan—because of robust issues, it had to limit its analysis to the first two years alone. However, this conclusion might be a result of the limitations of its assessment approach (ie, pooled OLS).[31]

A1.1.4 Models proposed for RIIO-ED1

Oxera understands that Frontier Economics has been developing TOTEX benchmarking models for the electricity DNOs.[32] Frontier Economics' preferred model uses data over five years (2006/07–2010/11), with TOTEX as the cost measure, and number of customers, peak capacity, population density and national wage index as explanatory variables. A time trend

[28] See, for example, surveys in Farsi, M., Fetz, A., and Massimo F. (2007), 'Benchmarking and Regulation in the Electricity Distribution Sector', CEPE working paper; and Jamsb, T. and Pollitt, M. (2001), 'Benchmarking and regulation: international electricity experience', Utilities Policy (9).
[29] Ofgem (2009), 'Electricity Distribution Price Control Review: Final Proposals - Allowed revenue - Cost assessment appendix', December.
[30] Ofgem (2007), 'Electricity Gas Distribution Price Control Review Final Proposals Document – supplementary appendices', December.
[31] Ofgem (2010), 'Consultation on strategy for the next gas distribution price control - Supplementary Annex - RIIO-GD1 Tools for cost assessment', December.
[32] Slides and data used by Frontier Economics were provided by ENWL.

is also included in the model to control for movement in costs over time, as a proxy to measure the technological change in the industry over the period. The estimation is done using a statistical model (a random-effects model), which attempts to estimate a company-specific component that is assumed to be invariant over time. This company-specific component is taken to be the measure of efficiency.

Frontier Economics has stated that its model provides sound economic interpretation and has reasonable statistical properties, while further work is required in some areas (in particular, on the real wages coefficient, positive sign on time trend and negative sign on density).

For RIIO-ED1, Ofgem has proposed using econometric analysis as part of its toolkit to determine the DNOs' relative efficiency. Despite this, Oxera understands that Ofgem is currently minded to use pooled OLS to estimate the efficiency levels of the DNOs unless there is a rationale for opting for approaches that control for company-specific factors in modelling.

A1.2 Ofwat's approach

For water services, Ofwat estimates econometric cost models for different groups of costs, such as resource and treatment or business support costs, using data for a single year using OLS. The results from the functional models are then aggregated to the overall OPEX level to obtain an overall efficiency target. With 18 companies providing water services, Ofwat compares costs at a company level. For the ten sewerage services companies, it undertakes econometric modelling at a sub-company level (eg, treatment works), or compares unit costs at a company level.[33] Ofwat used to assess CAPEX (normalised over several years) using OLS, but because of robustness issues, it has abandoned this method in favour of a unit cost approach in the previous two price reviews. Unlike Ofgem, Ofwat undertakes cost modelling on the unadjusted costs and make adjustments for company-specific factors to the estimated efficiency gaps in determining the overall efficiency target. Ofwat's criticism of modelling on adjusted cost (as Ofgem does) has been that since the size of these cost adjustments are often difficult to quantify, an attempt to correct modelled cost may introduce additional noise to the data and, instead of increasing the accuracy of the assessment, it may actually be reduced.[34]

For the current price review, PR14, Ofwat appears to have moved away from its use of cross-sectional data and is considering panel modelling approaches such as a random-effects model and panel SFA models for cost assessment.[35]

A1.3 The ORR's approach

The ORR used SFA to assess the efficiency of the rail network operator, Network Rail, using data from international comparators, and performed cross-checks with results from DEA and COLS. The ORR used an 11-year panel dataset of international rail network operators (12 other European operators) against which to benchmark Network Rail. For the ongoing review for the period 2013–18, the ORR is proposing to use advanced SFA models to assess Network Rail's efficiency,[36] based on recommendations made by Oxera (2009).[37]

[33] Ofwat (2009), 'Relative efficiency assessment 2008-09 – supporting information', December.

[34] See slide 3 of: http://www.ofwat.gov.uk/legacy/aptrix/ofwat/publish.nsf/AttachmentsByTitle/ms_part2110707.pdf/$FILE/ms_part2110707.pdf. Accessed on February 5th 2013.

[35] See CEPA (2013), 'PR14 Cost Assessment', a report for Ofwat, January.

[36] Office of Rail Regulation (2011), 'Establishing Network Rail's efficient expenditure', July, pp. 28–30.

[37] Oxera (2009), 'Recommendations on how to model efficiency for future price reviews', November.

A1.4 Ofcom's approach

Ofcom has used panel SFA to assess the TOTEX of BT's Openreach activities. Using publicly available information, BT was compared against the US local exchange carriers (LECs) by normalising BT's data to be comparable with that of the US companies.[38]

[38] For more information, see NERA (2008), 'The comparative efficiency of BT Openreach', a report for Ofcom, March.

Park Central
40/41 Park End Street
Oxford OX1 1JD
United Kingdom

Tel: +44 (0) 1865 253 000
Fax: +44 (0) 1865 251 172

Stephanie Square Centre
Avenue Louise 65, Box 11
1050 Brussels
Belgium

Tel: +32 (0) 2 535 7878
Fax: +32 (0) 2 535 7770

200 Aldersgate
14th Floor
London EC1A 4HD
United Kingdom

Tel: +44 (0) 20 7776 6600
Fax: +44 (0) 20 7776 6601

www.oxera.com