COST ASSESSMENT – USE OF PANEL AND SUB-COMPANY DATA
OFWAT

May 2011

REPORT

Prepared by:

Cambridge Economic Policy Associates Ltd

In association with

Mott MacDonald Ltd

CONTENTS

IMPORTANT NOTICE

This report has been commissioned by Ofwat. However, the views expressed are those of the consortium alone. The consortium members accept no liability for use of this report or any information contained therein by any third party.

Glossary of terms

Allocative efficiency. Allocative efficiency refers to the efficiency with which companies are allocating inputs (resources). A company will be allocatively efficient if no reallocation of inputs would produce a greater benefit for its customers or society.

Corrected ordinary least squares (COLS). Please refer to Ordinary least squares. COLS follows the same statistical technique as OLS (i.e. estimating a line of best fit by minimising the sum of squared errors), however the 'average' line is shifted towards the best performing company.

Data envelopment analysis (DEA). A quantitative non-parametric technique that optimises the number of inputs required for a particular output and vice versa. It does not require assumptions on the functional form, but it also does not allow statistical testing on the significance of explanatory variables.

Degrees of freedom. This term refers to the total number of observations in the sample less the number of independent constraints or restrictions (i.e., variables).

Efficiency change. Efficiency change refers to the 'catch-up' component of a company's productivity. The 'change' is calculated in relation to comparable 'frontier' companies or a benchmark (e.g. industry upper quartile efficiency performance).

Functional analysis. A form of sub-activity analysis focused on cross-cutting activities like power costs.

Menu regulation. Menu regulation is a form of regulation where regulated companies are no longer presented with a 'take it or appeal it' regulatory offer regarding the allowed level of expenditure, but are instead given a range of options from which to choose.

Non-parametric index numbers. Non-parametric index numbers refers to a method of estimating changes in efficiency, technical and productivity over time and across different comparators.

Ordinary least squares (OLS). OLS is a method by which linear regression analysis seeks to derive a relationship between company performance and characteristics of the production process. This method is used when companies have relatively similar costs. Using available information to estimate a line of best fit (by minimising the sum of squared errors) the average cost or production function is calculated and companies are benchmarked against this.

Panel data set. A panel data set refers to a data set containing cross-section observations collected for a number of years e.g. 10 companies are surveyed over 10 years. A panel data set can be created at any given level of aggregation (e.g. company, sub-company, sub-activity).

Pooled data specification. A pooled data specification (or approach) refers to an approach to modelling panel data that does not take into account any relationships between the observations either in a cross-section or over time.

Productivity change. Productivity change captures the combined changes resulting from: technical change; efficiency change, allocative efficiency and economies of scale.

Stochastic frontier analysis (SFA). Stochastic frontier analysis is a method of econometric modelling which explicitly tries to model 'noise' in the data measurement, thereby allowing the inefficiency component to be identified separately.

Sub-activity. Sub-activity refers to a specific activity within an identified vertical segment of a WoC or WaSC , for example sewage treatment plants.

Sub-company. Sub-company refers to a level below a company's total output at which outputs and input costs are identified separately. Sub-company refers to an entire vertical segment within a WoC or WaSC, for example water treatment (as defined by Ofwat's accounting separation guidelines). However, when referring to sub-company cost modelling we include sub-activities including functional analysis.

Technical change. Is the 'innovation' component of productivity and captures the shift in an industry's efficiency frontier.

WaSC Water and sewerage company.

WoC. Water only company.

EXECUTIVE SUMMARY

This report forms part of Ofwat's "Future Price Limits" project, and examines the extent to which Ofwat could make use of panel and sub-activity data in assessing the performance of companies. *Panel* data refers to the use of more than one year's data in assessing performance, be it backward looking historical information and/or forward looking forecasts. *Sub-company* data is more difficult to define since almost all information is to some degree sub-company, if only insofar as it relates to water or sewerage. For the purposes of this report, we have interpreted two separate aspects which are defined as:

- *sub-company* refers to the nine vertical segments that Ofwat has used for the purpose of defining accounting separation; and
- *sub-activity* means *below* the level of the nine vertical segments.

So data on water resources by zone, sewage treatment by plant or retail costs for a particular class of customer are all examples of sub-activity data. Analysis of segmental costs, such as IT or property or power costs, provide a different example of sub-activity data.

Cost modelling is widely used among regulators as a way of determining relative efficiency levels and thereby setting efficiency targets for natural monopolies. To date, Ofwat has made use of sub-company and sub-activity data, particularly for certain areas of operating expenditure, but has not routinely used panel data for comparative analysis. This report therefore examines the academic basis for such analyses and the extent to which other economic regulators have used these approaches. Ofgem, for example, relied heavily on panel, sub-company and sub-activity data in the 2009 price review of the Electricity Distribution Network Operators. Ofgem's recent RIIO review concluded that greater use should be made of forward looking panel data. ORR has used an 11 year panel dataset to compare Network Rail with 12 other European rail network operators.

This report also considers the context in which Ofwat will decide whether to use such data. This context is provided in part by the Cave Review, which identified several opportunities for increased competition in the industry, which will require new information for entrants and Ofwat. In isolation, the Cave Review can be expected to lead to greater use of panel, sub-company and sub-activity data. The Competition Commission has also recommended that Ofwat should consider using panel data. In addition, the Gray Report on Ofwat is due to be published soon and the indications are that it will recommend that regulation should become *less* intrusive and data intensive than it is at present. Ofwat is also carrying out its own project on regulatory compliance, one of the aims of which is to reduce the regulatory burden. There are therefore conflicting pressures on Ofwat, which reflect the fact that there are costs as well as benefits in adopting more sophisticated modelling approaches. The tests for change are therefore that they represent an improvement on the status quo and that the resulting benefits outweigh the costs. It has not been practical in this report to quantify costs versus benefits so the conclusions reached here rely on a subjective assessment of these issues.

Our overall conclusion in regard to panel data is that it would be beneficial for setting price limits at the next price review, PR14, and beyond. Although at PR14 greater care will be needed as the panel will be for only three or four years. Panel approaches are beneficial because they would increase the available observations (data points) used in the models and therefore improve the models statistical significance and/or allow more explanatory variables to be included. It would also enable trends over time to be measured at company and industry levels. Only limited or no additional information should be needed by Ofwat than at present, but their models would be able to use multi-year rather than single year data.

We consider that Ofwat should focus on backward looking data at PR14, but should also investigate the feasibility of using forward looking data, with the forward looking data taken from companies' PR14 business plans. There is currently a lack of practical evidence regarding the use of forward looking data and we recommend that Ofwat consider Ofgem's use of this data for its ongoing price determinations in gas distribution and energy transmission, RIIO-G1 and RIIO-T1. We believe that it would be unduly onerous to expect companies to reapportion pre-2010 costs across the nine businesses defined for the purposes of accounting separation. There will therefore be only a few years of historical data available for analysis at this level in PR14, but this should improve over time with regulatory commitment. We do however note that provided that appropriate tests for data consistency are carried out, there is the potential to employ aggregate company level panel data from before and after the introduction of the accounting separation guidelines.

It is also clear that sub-company data will be important in the future for price setting if greater competition is introduced, and segments of companies are excluded from a price cap or if local access prices need to be determined. Sub-company approaches have also been used by Ofwat in the past. Our overall conclusion in regard to sub-activity data, however, is that its use would be unduly onerous for setting price limits, with the exception of those small number of examples where Ofwat has needed to use sub-activity data for many years, such as for large sewage treatment plants and consequently some data is already being collected. We do, however, see a vital role for sub-activity data in facilitating competition. This is because most activities by entrants, be it in water abstraction, retail, sludge treatment or whatever, will be at a sub-activity level. In order to calculate access prices for an essential facility and to provide transparency of an incumbent's costs in potentially contestable markets, Ofwat will need to examine costs at a sub-activity and location specific level.

Any change to Ofwat's approach to modelling, including adopting the use of panel or sub-activity, raises concerns around both the consistency of results and the robustness of approach. This is in part because any change is likely to be linked to a different definition of costs or set of explanatory variables in the model. But it is also in part because there may be concerns about the regulatory appropriateness of the modelling. Of course, regulators are limited in what they are able to do and consequently must often trade-off the costs of data requests with theoretical requirements for modelling. However, these concerns do not mean that modelling should not change, rather that robustness checks should be introduced so that corrections for bias in the results can be introduced by the regulator. These robustness checks include running models at different levels of aggregation and using available statistical tests for bias in the results. If Ofwat's wider review of cost assessment and recovery recommends the continued use of economic cost modelling, then we recommend that Ofwat should consider adopting these sorts of checks alongside a move to panel data both at PR14 and beyond.

1. INTRODUCTION

Cambridge Economic Policy Associates (CEPA) in collaboration with Mott MacDonald and Dr. David Saal of Aston University have been retained by Ofwat to advise on the possible use of panel and sub-company data as part of the suite of performance measurement tools available to the regulator. Our team has involved individuals from a range of backgrounds, including economics, engineering, water and sewerage business and regulation. In addition, to ensure a robust approach to the project a key academic role has been taken by Dr David Saal.

This work is being undertaken as part of Ofwat's Future Price Limits project. Several other pieces of work are underway within this overall project, some of which are also linked to the cost assessment/performance measurement area. Consequently there are areas in this report where we note issues that need to be addressed but which are outside the narrow scope of our project but likely to be within-scope for other work that is underway. Clearly cost allocation and the implications for performance measurement are an important consideration within the accounting separation work underway in Ofwat.

This work has been undertaken through close collaboration with staff at Ofwat through a series of workshops. We appreciate the time and resources Ofwat has made available to work with our team but the views expressed in this report represent the thinking of our team and should not necessarily be attributed to Ofwat.

Clearly the industry is going through some major changes at the moment with significant uncertainty about the potential structure of the industry in five to 10 years. As such, rather than seek to provide a definitive answer as to whether panel and sub-company data is useful and would meet any cost-benefit requirements, we seek to determine, firstly, under what conditions these approaches might be appropriate, and, secondly, the issues which must be considered in applying these techniques. For the latter, in particular, we examine the issues on the basis of Ofwat's accounting separation guidelines. We believe that this framework provides a flexible foundation for Ofwat while still allowing robust and specific answers to be developed in the light of this report.

1.1. Structure of the report

This report is structured in three broad blocks as follows.

- Sections 2 to 4 provide key background to the question by considering the context of the project. This block includes a brief review of the economic theory underpinning performance/cost assessment and practical regulatory use of benchmarking;
- Sections 5 to 7 set out our approach to assessing the applicability of both panel data and sub-company data. This approach is then applied to one possible future industry structure. We primarily focus on the structure defined in the accounting separation guidelines that Ofwat has already introduced; and
- Sections 8 and 9 draw out the key lessons/observations from the analysis after undertaking a high-level qualitative cost benefit analysis of the approaches. We conclude by noting some possible further considerations.

A series of annexes support the report providing detail of existing regulatory and academic uses of both panel and sub-company data in regulated water, energy, post and rail industries as well as the implications of some other possible future industry structures.

In reading this report, it is important to bear in mind the terms of reference for this project. Our remit was to consider the case for panel and sub-company analysis, not the overall appropriate way that Ofwat should approach performance measurement. Of course, addressing our narrow issue without some consideration of the broader questions is not possible. We have, however, tried to limit the interaction with the broader questions to noting how and where these wider issues arise and we leave it to the consultants engaged by Ofwat to conduct the *Use of Comparator* study to answer those questions. Section 3 is the only section that does not follow as closely our attempt to limit ourselves to the direct question posed to us. This is due to the difficulties of separating the narrower and wider issues when considering alternative modelling approaches.

Aspects of this report are inevitably quite technical. In consequence, to make the report accessible to as broad an audience as possible, each section commences with a clear statement of the key themes and lessons in that section. This way we believe that the flow of the argument in the report should be accessible to all readers.

2. CONTEXT

Key themes and lessons

Up to now, Ofwat has placed significant emphasis on comparative performance measurement when determining prices at periodic reviews. Further, legislation currently requires the OFT to refer mergers (excluding smaller companies) to the Competition Commission (CC). The CC, if it deems a water merger has taken place, must assess the impact of the merger on Ofwat's ability to make comparisons between different water companies (including qualitative comparisons), demonstrating the central role that such performance measurement is seen to have within the regime.[1]

Ofwat developed models for the first review, PR94, which have formed the basis of such analysis at each periodic review from 1994 to 2009. However, as the amount of data available for analysis has increased over the years, it is now possible for Ofwat to adopt other modelling approaches to comparative analysis than those used up to now.

In addition, other stakeholders in the sector, including the CC, the Cave Review, companies and academics, have made recommendations that Ofwat could usefully adopt alternative approaches, especially panel data and sub-company analysis.

Changes are taking place within the sector with the introduction of the accounting separation guidelines, which require companies to provide data on nine different business units. In addition, the Government is considering potential future market reform, including possible industry structure changes. This means that the need for performance measurement and the type of analysis undertaken is changing, with the implication that the appropriate methods of comparative performance analysis may also need to change.

This section considers some of the rationale for considering Ofwat's approach to cost/performance measurement and the broader context for this project.

It should be noted that much of this section can also be seen as justification for other projects underway within Ofwat to consider not just the general way that performance measurement is used and undertaken but also the overall role of regulation. However, these elements are outside the scope of this project.

2.1. Ofwat's approach

Ofwat developed a detailed approach to performance measurement for the first price control, PR94, and while some changes have taken place since then, the basic approaches have remained the same. Ofwat's approach was to:

- model separately operating costs (opex) and capital costs (capex);
- assess capital efficiency by making comparisons of standard unit costs using the cost base tool;
- use single year cross-section econometric analysis of the companies to determine the relative opex efficiency of companies (Ofwat used to use an average cost over a number of years for capital maintenance analysis);

[1] Schedule 4ZA, Enterprise Act 2002.

- undertake the opex assessment at a range of levels, including functional ones (such as power costs) which are below the activity to which the price-cap is applied;
- use sub-activity data for large sewage treatment plants and sewerage areas and also use some unit cost models for sewerage opex; and
- use a time series based approach when considering the annual overall efficiency frontier shift.

The opex econometric models were developed by Ofwat in collaboration with academics in the early 1990s, and while they have been refined over time, the underlying approach and methodology has remained largely stable since the 1994 price review.

However, some of the refinements that have taken place since PR94 have included the removal of econometrics. For example, at PR09, Ofwat did not use the capital maintenance econometrics models. Instead it used the cost base tool together with the capital expenditure incentive scheme (CIS).

2.2. Issues

The approach to performance measurement that was put in place for PR94 was one that clearly responded to the constraints faced at that time, namely:

- limited historical data;
- a regulatory framework focused on water and sewerage level activities;
- a significant number of WoCs as well as the limited number of WaSCs; and
- the need to model a complex industry cost structure despite limited data.

The data available to Ofwat has been protected to some degree over the last 20 years as the merger rules established for the water sector required a clear link to the impact on the regulator's ability to undertake comparative assessment including comparative performance measurement. (As set out in the law and evidenced by the early Monopolies and Mergers Commission reviews.)[2]

What is clear is that since PR94:

- a more complete and consistent data series has been developed which makes additional approaches possible, as suggested by various academic papers (refer to Section 3.3 and Annex A.2, A.3 and A.4.1 for examples) that have employed Ofwat's regulatory data for panel analysis;
- possible trade-offs are lost through Ofwat's approach to modelling, such as between opex and capex and/or between functions or inputs (such as power and other opex);[3] and

[2] The 1989 Water Act.

[3] This issue is in part addressed through the discretionary way that Ofwat sets the final efficiency targets, but this causes other concerns with the regulatory regime.

- more complex/comprehensive models for costs can be considered if more data/observations are available.[4]

These issues are discussed further in sections 3 and 4 of this report.

What is clear, however, is that further consideration of the way in which Ofwat undertakes performance measurement is now highly relevant and has, in part, been captured in some of the more recent joint UKWIR/Ofwat studies.[5]

2.3. Other factors

Even if Ofwat's PR09 approach was "best practice" there are other factors that mean a reconsideration of the approach to performance measurement, and the role of panel and sub-company analysis, is needed.

2.3.1. Accounting separation

Ofwat has introduced accounting separation guidelines based around cost allocation for the companies from 2009/10 based on nine functions or business units within water and sewerage. These functions are discussed later in this report in some detail but broadly break into three businesses: water, sewerage and retail. Water and sewerage are then both broken into a further four activities. Ofwat is currently only asking companies to report their costs according to the nine functions set out in the accounting separation guidelines.

The introduction of reporting costs to the separate business units raises concerns about the existing performance measurement in two ways:

- it may cause a structural break in the data with both past data collected and between companies e.g. if the new accounting separation guidelines lead to a different cost allocation it may take a while for all companies within the sector to allocate costs in the same way; and
- the potential change in the way that regulation would be applied to the sector if there were to be a new separate retail control introduced at PR14. The final form that the regulatory structure will take is dependent on the proposed new legislation (which will stem from the as yet unreleased Department for Environment, Food and Rural Affairs (Defra) Water White Paper).

Accounting separation also raises additional issues related to the appropriate assessment of overall costs where cost interactions between various activities exist. This issue is discussed in detail in Section 3 of this report.

2.3.2. Cave Review

Developing from the accounting separation point but raising some more fundamental issues is the question of introducing market reform. Work on the question of where greater competition

[4] These issues/concerns have to be measured against the benefit of a relatively consistent approach having been adopted at the four price reviews that have taken place since privatisation in 1989.

[5] See for example: Ofwat/ UKWIR (2007) and UKWIR (2008).

could be introduced and the implications that this has for regulation of the water and sewerage industry is at the heart of Ofwat's work on Market Reform. The Cave Review clearly identified areas where greater competition was deemed to be possible – primarily around water abstraction and retail. It also recommended that the thresholds for merger referral be changed, something that could potentially allow a further consolidation in the industry.

Defra will publish later in 2011 a Water White Paper, aimed at providing clear direction and a policy framework for addressing future challenges in the water sector. However, the content of the Water White Paper is not yet known. While the paper can be expected to include a discussion on the issues around industry structure, whether the recommendations contained will include some degree of ownership/legal separation or just build on Ofwat's accounting separation guidelines is unknown.

In view of this, we examine the modelling issues in subsequent sections primarily on the basis of the accounting separation framework but also with some limited discussion of the implications from further separation.

2.3.3. Competition Commission

The CC has reviewed a number of mergers in the water sector. Cases considered by the CC have included referred mergers between South East Water and Mid Kent Water and Severn Trent/ Wessex Water and South West Water.[6] Strong recommendations have been made in some of the resulting CC reports for Ofwat to consider alternative approaches to performance benchmarking, including the use of panel data methods.

2.3.4. Other regulatory reforms

Ofgem completed its review of whether existing energy regulation in Britain was "fit for purpose" during 2010 – the RPI-X@20 review (then renamed RIIO for Revenue = Incentives + Innovation + Outputs). One of the areas considered in this review was performance measurement with a strong recommendation being made to focus on panel data and using forward looking information. This recommendation in part built on the last electricity distribution price determination (made in 2009) which used panel data analysis, although the RIIO focus on forward looking cost estimates was new. Forward looking data is to be drawn from highly detailed company business plans, submitted to Ofgem for the purpose of assessment and benchmarking. It is not yet clear the extent to which Ofgem are proposing to use forward looking data for performance assessment whether this is target setting or just consistency checking. However, Ofgem's current strategy for the first RIIO transmission price control (RIIO-T1) states that business plans will be reviewed in order to determine ongoing efficiency improvements.[7] For example, Ofgem states that transmission operators' (TOs) direct operating expenditure "will primarily be based on our assessment of the forecasts in their business plans".[8]

---

[6] The Severn Trent/Wessex Water and South West Water reviews were undertaken by the Monopolies and Mergers Commission.

[7] Ofgem (2011a).

[8] Ibid, p 3.

Other regulators around the world have used different approaches to performance measurement some of which include the use of panel and sub-company assessment. These are discussed further in Section 4 below.

2.4. Summary

It would be appropriate to consider whether panel and sub-company data could improve Ofwat's performance measurement tools, even if the potential changing industry structure, changing focus of regulation and lessons from other sectors (especially energy in Britain) were not also creating pressure for such an assessment. However, the development and implementation of accounting separation guidelines provide another reason why such an exercise is worth doing at this time.

3. COST MODELLING USING PANEL AND/ OR SUB-COMPANY DATA

Key themes and lessons

Our findings show that it would be possible for Ofwat to measure performance using panel and/ or sub-company data. However it is important that however Ofwat chooses to measure performance it considers the appropriate economic theory, and addresses any issues that arise.

From theory the following key themes and lessons arise for panel data and sub-company analysis.

Panel data

- Panel data has great potential to improve Ofwat's ability to carry out cost assessment in the water industry. This is because it can increase the available degrees of freedom which allows for improved model specifications, which could make regulatory cost assessment models more consistent with economic cost modelling approaches. This can allow for better modelling of differences between firms.
- If Ofwat were to use panel data, it would require better controls for intertemporal differences in firm characteristics, capital stocks, input prices, and possibly technical change. Given that there are a limited number of potential observations and the additional parameters will use up some of the additional degrees of freedom, we can see that panel data will by no means be a panacea for effective performance measurement.
- If Ofwat used panel data, then it would be able to use benchmarking techniques that better allow for opex-capex tradeoffs. This is because panel models have the potential to provide sufficient data to allow modelling of opex costs as a function of input prices, MEA capital stocks, and operating characteristics. Panel data therefore could facilitate the adoption of benchmarking approaches that result in better incentives for companies to improve their overall cost efficiency

Sub-company and sub-activity data

- There is no reason in principle why Ofwat could not use sub-company data modelling to model firm costs, provided there was sufficient regulatory commitment to data collection. Ofwat would need to collect appropriate data on outputs, inputs, input prices, capital stocks, and operating characteristics *at the sub-company/activity level.*
- Because of substantial geographic heterogeneity in water industry operating characteristics, sub-company data may allow a better understanding of the determinants of costs. However controlling for these operating characteristics may require additional degrees of freedom. Panel data may therefore be required to increase the number of observations and hence the degrees of freedom.
- Modelling with sub-company/activity data has the potential for substantial biases related to aggregation. For example, firms may be horizontally integrated (at the sub-activity level) because of the presence of horizontal integration benefits accrued from multiple plant operation. Ofwat should therefore apply aggregation techniques that take horizontal integration into account.
- Ofwat must also contend with sub-company/activity analysis issues related to vertical cost interactions if they influence costs.
- There are a number of potential benefits to using sub-company data, and they could be used to facilitate the development of efficient competition. However, effective benchmarking at sub-activity level may potentially be both infeasible and prohibitively expensive given the amount of data required.

Cost models

- There are a number of different methods that can be used for cost modelling purposes with sub-company, sub-activity or panel data. Commonly used approaches by regulators and academics include, econometrics, data envelopment analysis, and non-parametric index number approaches.
- Each approach has its own benefits and limitations. We recommend that Ofwat employ index number approaches as a readily applicable company level consistency checking tool, with which relative company performance measures derived with econometric or DEA methodologies can be cross checked.

3.1. Introduction

This section briefly considers the key concepts of economic theory of cost modelling that are relevant to the water and sewerage sectors and, given this context (and that of regulatory objectives), discusses the potential applicability of different approaches. This will allow us to highlight several generic modelling issues that are likely to influence model accuracy. Annex A builds on the themes presented below and provides, from an academic viewpoint, a more complete consideration of the economic theory of cost modelling and the potential implications for regulatory cost assessment.

3.2. Economic theory of cost modelling

The economic theory of costs rests on the assumption that firms aim to minimize the total cost of production given the production technology available to them, the outputs they produce, and the input prices they face.[9] It also highlights the need for fully flexible modelling approaches able to allow for a more complete specification of the underlying relationship between costs, outputs, input prices, and the scale of operations. As a result, economic cost modelling generally requires the specification of costs as a function of outputs and input prices, and the use of an empirical model that is sufficiently flexible to reflect how alternative methods of production influence a firm's costs. Fundamentally, such flexible models allow for the cost implications of important economic relationships to be captured in the empirical modelling.[10] To illustrate this, in this section we focus on the economic concepts of opex capex trade-off, cost interactions, scale, heterogeneity and technical change.

Opex capex trade-off

Water industry capital stocks are not fully flexible and adjust slowly over time. Nevertheless, differences both between companies and across time in these capital stocks have a significant influence on opex requirements.[11] The combination of: (i) a complex output production process; and (ii) input substitutability between opex and capex in the water industry suggests that high level total cost benchmarking approaches which are broadly consistent with economic total cost modelling are appropriate. This is particularly the case for longer panel data analysis where

---

[9] See, for example, Chambers (1988).

[10] For example, Stone & Webster (2004a) employ a fully flexible cost function specification which suggested increasing returns to scale for small WoCs, but larger WaSCs exhibited decreasing returns.

[11] See for example, Stone & Webster (2004b) which found that changes over time in water and sewerage capital stocks had an impact on opex requirements.

capital adjustment is more likely to approximate long term adjustments.[12] An alternative is the use of economic quasi-fixed capital stock modelling approaches that allow for both the relatively fixed nature of capital, while also allowing for the impact of capital stocks on opex requirements. This also has the potential of substantially improving the assessment of opex costs. Assessment of efficiency or performance by regulators will generally only focus on costs within management control. Either of these approaches would better allow for opex capex tradeoffs than Ofwat's current approach. Section A.3. further discusses the appropriateness of total cost benchmarking in Water and Sewerage cost modelling.

Cost interactions

Production of water and sewerage services involves a complex multiple output production process (with the potential for vertical cost interactions in water and sewerage service production).[13] Economic cost modelling theory and practice suggests that such multiple output production requires approaches that fully account for the potential cost interactions between these activities. It also suggests that, in the presence of such cost interactions, an aggregate cost assessment which allows for these multiple outputs is most appropriate. However, if the costs of two activities are economically separable, unbiased cost assessments of each activity can be obtained with separate assessment of each activity. In contrast, if vertical cost interactions exist there is strong potential for overall cost assessment estimates to be biased unless the separate assessments properly account for any cost interactions between other activities which influence the costs of the activities under assessment.

Similarly, if horizontal integration economies are obtained by aggregating treatment plants, water supply zones, and other sub-company/activity operations (at least to some optimal level of horizontal scale), assessment at a level below that of the company may not properly allow for any benefits of horizontal integration and will result in upwardly biased estimates of aggregate costs for larger companies if horizontal integration economies are present. In contrast, if diseconomies of horizontal integration are present, larger firms might have downwardly biased estimates of aggregate costs.

The above therefore suggests the potential need for a regulatory cost assessment process of identifying and testing for the presence of any relevant cost interactions between activities. This would include employing appropriate controls to account for such cost interactions and, if carrying out the cost assessment at a level where such interactions are internalized, using higher level assessments as consistency check of disaggregated assessment, and/or employing appropriate regulatory discretion to mitigate any potential assessment biases.

Scale

Given the possible cost interactions, sufficient model flexibility is required in order to properly assess the relationship between a firm's output scale and its costs. Fully flexible modelling techniques allow for variable returns to scale (i.e. constant, increasing and decreasing), and are

[12] Baumol, Panzar & Willig (1982) note that consideration of potential cost interactions between multiple outputs is necessary to properly assess costs.

[13] See Bottasso & Conti (2009b) and Saal, Arocena & Maziotis (2010) for evidence of vertical integration in the English & Welsh water and sewerage sector.

therefore able to result in estimates reflecting the likelihood that smaller firms would benefit more than larger firms from increases in their scale of operations. Work commissioned and published by Ofwat, suggested the presence of variable returns in the English & Welsh water industry, with the smallest firms characterized by increasing returns to scale and larger firms in the industry characterised with decreasing returns to scale.[14]

While, it can be argued that regulators should assess performance based on the costs of a firm operating at the optimal scale, and thereby penalise firms that are below and above this scale, it may not necessarily be the case that regulated utilities are fully in control of their scale as they must serve their licensed operating area and its population. Thus, utilities should be assessed with a methodology that can consistently capture the cost implications of their scale of operations.

Heterogeneity

The costs associated with water and sewerage services are substantially influenced by heterogeneity in operating characteristics. These include differences in topography, settlement patterns, water source type, quality, location, and required water and sewerage treatment technologies, etc. These characteristics can also vary substantially within reasonably small geographic distances; they also can, and do, change considerably over time. Including improved controls for such heterogeneity directly within cost assessment models should improve the assessment. Moreover, such direct controls would potentially give Ofwat the ability to allow the model to determine whether an exogenous company specific cost driver is in fact a significant determinant of cost variation between firms.

The potential for input substitution and the incentive for managers to minimise costs given the input prices they actually face, suggests the need for careful controls of input price variation between firms, particularly when there exists considerable cross-sectional variation in relative input prices.[15]

Technical change

Ofwat's regulatory practice of identifying continuing efficiency improvements is consistent with an assumption that technical change (or equivalently frontier shift), does occur in the water industry.

The appropriate modelling of technical change in a fully specified water industry cost model is likely to require a sophisticated modelling approach as, over time, technical change may not influence in a proportional way the relationship between all inputs, outputs, and costs. This implies that approaches such as including time trends or time dummies, which allow only for neutral technical change, will insufficiently capture changes in cost functions over time. Instead, reliable models are likely to require fuller parameter flexibility that better reflect the impact of technology on costs over time.

---

[14] Stone & Webster (2004a).

[15] Saal & Parker (2000) found that there was evidence to suggest considerable gains in WaSC cost productivity attributable to the substitution of other inputs for capital. The findings in Maziotis, Saal & Thanassoulis (2011) also support this.

3.3. Cost modelling techniques

This section summarises key features, strengths and limitations of different techniques that can be used for performance assessment. Annex A and Section 7 provide a more detailed discussion of the characteristics of these approaches and their potential application in regulatory cost assessment. We focus on three alternative approaches to estimating firm performance, which are the econometric (including ordinary least squares, corrected least squares, and stochastic frontier analysis), data envelopment analysis (DEA) and non-parametric Total Factor Productivity (TFP) index number approaches.

3.3.1. Econometric techniques

Econometric approaches are regression-based approaches that rely on the specification of a production or cost function, thereby providing a fitted functional form capturing the underlying economic determinants of production and costs, such as the relative magnitude of input usage on outputs, or the impact of increased outputs and input prices on costs. These approaches can be further classified into approaches that attempt to explain the average relationship between costs and outputs, represented in this discussion by ordinary least squares, and those that are intended to identify the best practice, or frontier relationship between costs and outputs, represented here by corrected ordinary least squares, and stochastic frontier analysis.

Ordinary least squares (OLS)

Ordinary least squares is a type of average response model. Such models estimate a line of 'best fit' to observed data points by minimising the sum of the squared deviations of the observations from the fitted line. An average response model therefore simply determines the expected or mean relationship between costs and a given level of outputs. Regulators have taken the mean fitted line to represent the costs that a company of *average* efficiency would incur. Such methods, when correctly specified and with sufficient numbers of observations available, provide the potential for extremely robust estimates of a firm's relative efficiency.[16] However, by statistical assumption, in an average response model any deviations from the mean are assumed to be noise.

We also note that the principal potential regulatory weakness of average response approaches is that they do not allow for the estimation of frontier costs, and hence do not provide a direct estimation of inefficiency that is theoretically consistent with the efficiency catch up component of a regulatory X factor. Nevertheless, they are well established and particularly when coupled with standard econometric approaches to modelling with panel data, they have the potential to provide effective estimates of the underlying relationship between outputs and costs, and the determinants of trends in those costs.

---

[16] It is well established that average response techniques work well with pooled cross-section and time-series data. See for example, Baltagi (2005). The use of pooled data increases the available degrees of freedom and has been shown to improve statistical robustness. This approach was suggested in Kumbhakar & Horncastle (2010).

Corrected ordinary least squares (COLS) and stochastic frontier analysis (SFA)

COLS and SFA are both types of frontier econometric techniques which are designed to provide an estimate of absolute inefficiency. Nevertheless, there are considerable and important differences between deterministic approaches such as COLS, which do not allow for statistical noise, and SFA, which does attempt to separate measures of inefficiency from measurement error (or noise).[17]

COLS has generally perceived benefits related to simplicity and transparency. The COLS approach relies entirely on an underlying estimated average response OLS model where the line of best fit is "corrected" by shifting it in parallel until it passes through the 'frontier' data point, forming a frontier model.[18] COLS is based on the assumption that the estimated OLS residual, the distance a company is from the line of best fit, can be interpreted entirely as a measure of inefficiency. However, it is likely that the residual will contain some amount of measurement 'noise'.

The SFA approach attempts to address this issue by explicitly separately estimating inefficiency and random error/noise. However, this advantage does come at the cost of a more complex econometric specification and it also requires relatively strong assumptions in regard to the statistical distribution of inefficiency. These assumptions are needed to identify the decomposition of the residual into inefficiency and random error and, if these assumptions are not valid, the resulting estimates and decomposition will be biased. SFA also requires a minimum of at least two more degrees of freedom than an average response model in order to separate noise and inefficiency in the estimated model. The period covered, available variables, and data consistency of Ofwat's accounting separation databases will therefore be the fundamental determinant of the feasibility of applying SFA in future cost assessment.[19] Even if it is used, it is important that SFA results are compared with the estimates from other and less complex methods to test their robustness.

A general advantage of all econometric approaches relative to DEA and non-parametric index number approaches is the potential to conduct hypothesis tests between alternative model specifications. This allows for direct comparisons between models and also, provides greater statistical confidence in the reliability of the chosen model. A general disadvantage, though, is the need to specify a functional form for the cost function.

---

[17] See Greene (2008) for a detailed discussion focussed on the cross-sectional estimation of deterministic and frontier efficiency econometric models.

[18] Regulators do not necessarily use the most "extreme" frontier company as the frontier data point (see, for example, the discussion of Ofwat's and Ofgem's use of COLS in Section 4).

[19] There are a number of academic papers that support the use of SFA, for example Kumbhakar & Horncastle (2010), and ORR also uses SFA as its main approach. However, an UKWIR (2007) study found that SFA was not appropriate for the Water industry.

3.3.2. Data envelopment analysis (DEA) technique

Traditional DEA models rely on linear programming techniques to construct a non-parametric frontier or envelope around the data, thereby providing a fitted production frontier detailing the relationship between a firm's outputs and efficient input quantities. Firms using inputs in excess of those indicated on the DEA frontier are therefore defined as productively or technically inefficient, with the distance from the frontier indicating the level of inefficiency

Appropriate estimation of cost or allocative efficiency in the DEA framework requires data on both input prices and quantities, as well as a specification of additional constraints (e.g. convexity, variables returns to scale) to the DEA programme. The latter have the effect of identifying the unique cost minimising point on the estimated production frontier.[20] Thus, while it is common for many researchers to specify DEA "cost models" by simply employing opex or other measures of cost as an input, this approach is consistent with neither economic nor DEA theory. Thus, as with econometric approaches, it is likely that regulatory DEA-based cost assessments can be substantially improved with appropriate controls for input prices, as managers will choose to use alternative combinations of inputs if the relative prices of inputs vary substantially across companies or time.

The principle advantages of DEA relate to its relative simplicity and the intuitive and fully flexible relationship between actual data points and the identified firms (peers) that define the production frontier.

The principle disadvantage of the traditional DEA approach is its failure to allow for random error and noise, as can be done in econometric modelling. In addition, while DEA estimation strategies do allow for the inclusion of operating characteristics as a determinant of costs, it is our opinion that traditional DEA models are far less capable of controlling for heterogeneity between firms or at the sub-company/activity level than econometric approaches.

As with econometric approaches, extending DEA techniques to panel data requires consideration of the possible presence of technical change, and if present an appropriate modelling approach to allow for it is required. One possible approach is to estimate a distinct DEA frontier for each year in the data, thereby allowing for the determination of efficient frontier costs in each year, as well as technical and efficiency change between years. However, this approach amounts to the estimation of separate cross-sectional DEA frontiers in each year, and would not help to overcome the issue of estimating costs if the number of cross-sectional observations was limited. We therefore believe there is unlikely to be any strong potential for developing this approach, or even any simple cross-sectional DEA assessment with Ofwat's regulatory data at firm level.

An alternative approach is sequential DEA modelling of panel data. This approach estimates a series of yearly frontiers, which are estimated for each year by sequentially adding each new year's data to a common data pool and re-estimating the frontier. This approach allows the identification of efficient costs in a given year, and hence allows for the separate identification of both technical change and efficiency change. However, if cross-sectional data is sparse, it would

---

[20] See for example Thanassoulis (2001).

still require the pooling of several years' data before the year of assessment to allow the estimation of frontier costs in that year.[21]

We conclude that while it would be worthwhile to further explore the potential application of DEA approaches, on balance, econometric approaches are potentially superior for direct regulatory benchmarking in the England & Wales context.[22]

3.3.3. Non-parametric index number techniques

Non-parametric TFP index numbers essentially use input and output prices to provide information on the relative costs of inputs and the relative value of outputs. This allows the construction of aggregate indices of outputs and inputs, from which TFP indices can be constructed.[23] The clearest benefit of this approach is that it can be easily employed to provide relative performance estimates with limited data as only data on input and output quantities, total costs, total revenues and prices are required. The approach requires a minimum of only two firms to allow meaningful cross-sectional performance comparisons between firms, and similarly requires only two years of data for a single firm to consistently assess trends in performance. Moreover, cross-sectional indices provide easily understandable measures of relative productivity between firms, which are inversely proportional to the potential efficiency catch up of the laggard firm relative to the best practice firm.

This ready applicability does however come with some shortcomings. Firstly, non-parametric index number approaches require the availability of reasonable proxies for output prices as well as input prices, thereby making their application infeasible in the absence of output prices. Secondly, the approach is relatively inflexible with regard to controlling for differences in operating characteristics and output quality between firms, thereby making appropriate controls for cross firm heterogeneity relatively difficult to implement. Thirdly, the approach assumes the presence of constant returns to scale, which is a significant disadvantage given our above discussion with regard to the appropriateness of allowing for fully variable returns to scale when conducting cross company comparison.

Nevertheless, despite these potential shortcomings, there is sufficient potential to apply index number techniques at company level, and to gauge relative performance trends between firms and across time. We therefore suggest, that particularly because of their applicability with an extremely low number of observations, Ofwat could potentially employ index number approaches as a readily applicable company level consistency checking tool, with which relative company performance measures derived with econometric or DEA methodologies can be cross checked. Further, it is possible that future industry structures could improve the availability of output prices at different points in the value chain. Consequently index numbers may become more easily applied. This point is discussed further in Section 7.

---

[21] For example, if only 10 cross-sectional observations were available, and the analyst believed that a minimum of 30 observations were required to adequately model firm costs, a frontier could only be specified for the third and any subsequent years.

[22] Bogetoft & Otto (2011).

[23] Chapter 4 of Coelli, et al. (2005) provides a general introduction to the economic theory of index numbers and productivity measurement as does Hackman (2008).

3.4. The use of panel data in water and sewerage cost modelling

A panel data set includes both a cross-sectional and time dimension and therefore can include multiple observations of firms in different years. A balanced panel data set is one in which all the cross-sectional units (companies) are captured in each time period and an unbalanced data set is one in which the cross-sectional units are not necessarily captured in each time period. There is however, no clear advantage from an econometric perspective to having a balanced or unbalanced panel.

With panel data, there is potential for both cross-sectional and intertemporal heterogeneity. These cross-sectional differences can, for example, come from heterogeneity between the underlying operating environment and costs of a firm. Similarly, intertemporal variation across a firm's observations can, for example, be attributable to technical change. Thus, panel data modelling generally requires controls for such cross-sectional and intertemporal heterogeneity.

A pooled model simply pools the available panel observations in the database and estimates a standard cross-sectional model without allowing for differences in parameters or model specification to allow for heterogeneity.[24]

The employment of panel data has the potential to improve cost assessment in the water industry by increasing the available degrees of freedom and thereby allowing improved model specifications which will allow for better modelling of firm heterogeneity. Heterogeneity will exist in both the cross-sectional and intertemporal dimension. Operating characteristics such as customer demographics, water source reliability and quality, mandated water and sewage treatment levels, operational quality of networks (leakage, pressure, etc), etc, will vary across time, and such variation will increase with the length of the panel data set.[25]

Beyond the general increase in degrees of freedom provided by panel data, there is also clear demonstrated potential to allow direct estimation of cost productivity trends, as well as decomposition of these trends into technical change, efficiency change, and other potential components. This implies the potential directly to estimate past achievements in overall productivity growth, efficiency change, and frontier shift, which could be employed to inform Ofwat with regard to a firm's potential future "X" factor cost savings. It could also make regulatory cost assessment models more consistent with economic cost modelling approaches. Hence, the regulatory use of panel data could provide Ofwat with the necessary degrees of freedom required to model more closely the relationship between scale and firm costs.

Whenever panel datasets of more than trivial length are employed, there is the potential for technical change. Thus, with panel data, models should make allowances for technical change, thereby allowing for the potential that any or all parameters may change over time.[26] However, even if only a short dataset is available pooled data may allow more robust estimates than Ofwat's existing model(s). In consequence, the following estimation strategy is appropriate with pooled data. First, it should be tested for technical change and, if technical change is not present then a straight pooling approach is appropriate; second, if technical change is neutral than a time

[24] See Section A.7.1 for further discussion on pooled specifications.
[25] Coelli (1999) provides a simple example of how such heterogeneity can influence estimates of efficiency.
[26] As suggested by Botasso & Conti (2009b), and Stone & Webster (2004b).

variable or year dummies should be included to control for this; but, third, if technical change is present but is either output or input biased then a more sophisticated model specification is needed to allow for that.

The statements above particularly apply to the econometric approaches, but past academic work has shown that panel data can be employed with econometric, DEA, and index number approaches.[27]

The use of panel data also opens up the potential for benchmarking techniques that better allow for opex-capex tradeoffs. This is because panel models have the potential to provide sufficient data to allow for high level total cost benchmarking, or alternatively modelling of opex costs as a function of input prices, MEA capital stocks, and operating characteristics. Panel data therefore could facilitate the adoption of benchmarking approaches that result in better incentives for companies to improve their overall cost efficiency. However, any movement to the use of panel data assessment will require substantially greater consideration of the impact of input prices on assessed costs. This is because it is implausible to assume that input prices will not change across time, and that relative input prices, for either the inputs used by a single firm, or those reflecting price differences across firms will stay constant across time. For example, the relative price of labour to capital tends to increase over time, in tandem with a steady substitution of labour with capital in most production processes where this is feasible.

Classical panel econometric models such as the random and fixed effects specifications allow for a firm specific error term in addition to the standard white noise regression residual, and therefore provide a form of control for firm specific heterogeneity. However, there is the risk that this heterogeneity includes persistent inefficiency as well as differences in operating characteristics. Thus, recent developments in panel techniques such as Greene's (2005) true random fixed effects model attempts separately to estimate inefficiency from firm specific heterogeneity.[28] However, as with SFA, this comes at the cost of greater complexity and a reliance on the statistical assumptions used to identify the different effects.

Finally, it is worth noting that as appropriate controls for cross-sectional and intertemporal variation in operating characteristics requires the inclusion of further explanatory variables in assessed models, the additional degrees of freedom provided by panel data, may not increase proportionately with the length of the panel data set. This is most likely to be a significant issue with short panels e.g. those at the sub-company or sub-activity level.

3.5. The use of sub-company and sub-activity data in water and sewerage cost modelling

There is no reason in principle that sub-company/activity data modelling could not be employed by Ofwat to model firm costs, provided there was sufficient regulatory commitment to data collection. This would require appropriate data collection on outputs, inputs, input prices, capital stocks, and operating characteristics at the sub-company/activity level. However, the data

[27] For example, Stone & Webster (2004a); Bottasso & Conti (2009a); Saal, Arocena, & Maaziotis (2010); Green (2005); Kumbhakar & Horncastle (2010); and Lawrence (2004).

[28] Application of this approach by Saal, Parker, & Weyman-Jones (2007) to a WaSC database, reveals that even after allowing for a fixed group specific heterogeneity term, firm specific operating characteristics still significantly influence input requirements in the water sector.

requirements for pooled or panel data sub-company/activity modelling are significant, including maintaining consistency of classifications.

Given substantial geographic heterogeneity in water industry operating characteristics, sub-company/activity data may well allow better understanding of the determinants of costs. Allowing for this heterogeneity through the use of sub-company/activity performance models could lead to improved efficiency estimates. Modelling at this level may be particularly beneficial if policies are implemented which are meant to improve performance by facilitating competition. This is because high heterogeneity implies the need for effective access pricing and the development of local access prices over time.

Sub-company/activity data analysis may also be supported by the evidence of company internal operations based around multiple operational areas – effectively sub-company/activity information. Many companies operate on the basis of more than one internal operational area, possibly in part reflecting the company's views about the optimal size of operations. If consistent data is available on the basis of these operational areas then this would support sub-company/activity analysis.

Modelling with sub-company/activity data, however, raises issues around the potential for biases related to aggregation. Companies may be horizontally integrated because of the benefits accrued from multiple plant operation.[29] If this is the case then appropriate aggregation techniques will need to be employed. Sub-company/activity analysis must also contend with issues related to vertical cost interactions if these influence costs. In addition, sufficient model flexibility may also be required in order to properly assess the relationship between a company's output scale and its costs. Thus, utilities should be assessed with a methodology that can consistently capture the cost implications of their scale of operations.[30]

3.6. Key considerations for the use of sub-company/activity and panel data

Table 3.1 below summarises the key things we believe that Ofwat needs to consider when deciding whether or not it should use sub-company/activity and/or panel data. We pick these issues up later in this report when we discuss the feasibility and appropriateness of sub-company/activity and panel data.

[29] Although, it should be noted that companies existing structure may reflect its historical base rather than an optimal structure.

[30] It should be noted that under some models of competition the scale of operation could change.

Table 3.1: Cost modelling key considerations summary table

Issue	**Concern and implications**	**Possible solutions**
Specification and application of underlying economic theory	*Input costs* – changing relative and intertemporal input costs over time can make model estimation difficult and introduce biases if specific input costs change relative to RPI (i.e. changes in relative input prices will affect firms cost minimisation decisions). *Heterogeneity* – companies (and sub-company segments) have different operating characteristics (e.g. source of water, urban density, etc) owing in part to different mixes of inputs and also different outputs (such as changing quality levels over time). *Opex capex substitution* – the shift towards greater capex and the impact this has on the estimate of performance when only one element, opex, is evaluated by Ofwat. *Theoretical underpinnings of the model* – is the model specification appropriate?	*Input costs* – Ofwat could address this by making real price effect ('RPE') adjustments,[31] and consequently it may not be an issue. Ofwat could make these adjustments to specific costs feeding into the models (i.e. a regional cost adjustment). Even if Ofwat did not make these adjustments, many models ignore this concern, and Ofwat could take account of this issue in the way it used the results of the model. *Heterogeneity* – If Ofwat increased the number of explanatory variables and ensured that aspects like quality were adequately captured in the outputs then this would capture at least some aspects of heterogeneity. (Increasing the number of variables will require a greater number of observations.) *Opex capex substitution* – Ofwat could overcome this concern by using models that capture both types of cost, such as totex or opex with an MEAV capital value. Totex has additional definitional issues (i.e. does it include capex, or depreciation and return on planned capex?) which may make it a less preferred solution.[32] *Theoretical underpinnings of the model* –Ofwat should sense check the cost function to ensure that it conforms to economic principles. Ofwat should also use statistical checks of the model(s) specification.
Technical change	Technical change can lead to shifts in the production frontier which may be captured as efficiency improvements. Technical change should be controlled for in order to determine the efficiency catch-up that may be achievable by a company/ sub-company/sub-activity.	The way that Ofwat undertakes the performance assessment is important for addressing this concern. Ofwat could: • use sequential DEA structure; or • include additional parameters within an econometric panel data approach to control for technical change.
Cost interactions	If the performance assessment is undertaken at a level where vertical cost interactions are not internalized then biased estimates may be derived.	Ofwat could assess vertical cost interactions by undertaking performance assessment at a higher level within the structure so that it can produce an unbiased estimate. It could then use this to inform its decision about

[31] Real price effect adjustments refer to making adjustments for observed differences between input price indices, i.e. an adjustment to an input if the input's price was growing faster (slower) than RPI.

[32] For example see Frontier (2010), a report commissioned by Ofgem as part of RPI-X@20 which investigated the use of totex in future price controls.

Issue	Concern and implications	Possible solutions
	If the performance assessment is undertaken at a sub-company/activity level where horizontal cost interactions are not internalized then biased estimates may be derived.	targets at a lower level or more generally when applying regulatory discretion to setting the target. Alternatively it could include controls in the modelling to allow for cost interactions (e.g. in a similar way to controlling for heterogeneity). Ofwat could assess horizontal cost interactions by undertaking performance assessment at a more aggregated level within the industry segment to produce an unbiased estimate. It could then use this result to inform a decision about targets at a lower level or more generally when applying regulatory discretion to setting the target.
Scale	If different companies are at different points relative to the optimal size for an industry, say some have economies of scale while others have diseconomies of scale then the model specification becomes very important otherwise biased results may be found.	Ofwat could: • use a "more flexible " model definition that better allows for variable returns to scale within the industry; and/or • use sub-company/activity data where appropriate – for example, if a water company has several separate operating units that are closer to the optimal scale than the whole business and which are treated like separate/independent operations, then measuring performance for those units rather than the aggregated business would be appropriate. These solutions mean Ofwat would need a good understanding of the economies of scale within the industry.
Length and breadth of panel	A sufficient number of observations is needed, this is a function of both the number of years of data as well as the number of cross-sectional observations	While ideally a long (and wide) panel would be available to Ofwat, it could also use a short panel. It would need to consider: • Pooled panel data specifications – this has other concerns relating to technical change but may be an acceptable approach in the short-term; • Testing for the presence of technical change and controlling for it if evidence for it is found, which is likely in a longer and well specified panel model; and • Use of sub-company/activity data to enhance the number of observations.

4. REGULATORY APPROACH TO COST MODELLING

Key themes and lessons

UK based regulators have historically adopted a building block approach to setting base revenue allowances. They have focused on determining efficiency targets for each of the major cost areas separately and then using *ex post* incentives to ensure that the capex opex trade-off that regulated companies face is taken into account. This differs from the approaches put forward by numerous academics (see Annex A) which emphasise the importance of taking account of this trade-off in the modelling *either:* (a) by controlling for capital in the opex modelling; *or* (b) by focusing on total economic cost modelling.

Regulators have to date avoided more complicated models for, but not limited to, the following reasons:

- concerns about the data quality;
- concerns about the complexity of the modelling and the reproducibility of the models;
- difficulties with developing accurate and consistent (across companies and over time) measures of capital stocks;
- the costs of greater data requirements and transparency of the modelling; and
- concerns that the benefits may be marginal.

Regulators have, in general, been less willing to adopt more sophisticated statistical techniques for their modelling, but there has been an increasing use of panel data and, where appropriate, sub-company data. However, in the case of sub-company data this has primarily been in assessing different areas of costs rather than separate vertical integrated segments. Ofgem has successfully used panel data across opex and capex, and as part of the electricity distribution price control review to expand the number of observations. It is also looking to extend the use of panel data by benchmarking forecast data points for the transmission and gas distribution price control reviews. Postcomm has made significant use of sub-company data in its analysis and ORR has used panel data (from international operators) to benchmark Network Rail against.

All regulators note the importance of conducting top-down benchmarking as well as bottom-up analysis given that bottom-up analysis may not capture the full scope for overall efficiency improvements. For example, not all mechanisms available to a company for increasing efficiency can be foreseen at the start of the price control. Bottom-up analysis focuses on identified initiatives and therefore may underestimate the scope for forward efficiency gains.

4.1. Introduction

This section briefly reviews Ofwat's PR09 methodology as well as that of other regulators. In it we consider how lessons from the economic cost modelling theory, as well as regulatory modelling considerations influence the appropriateness of these models.

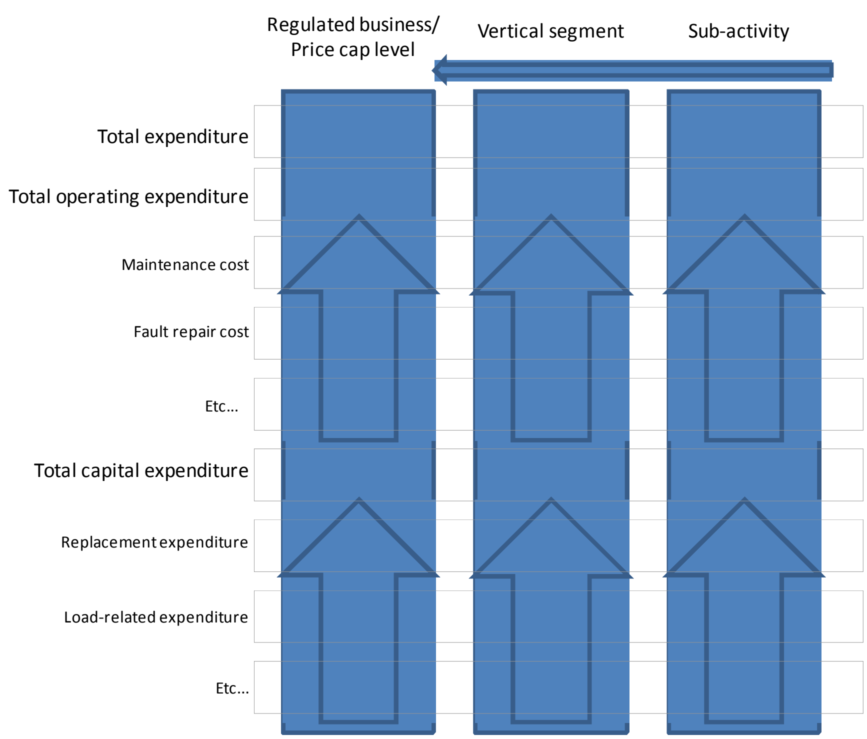
4.2. Building blocks approach

4.2.1. Overview

Historically, utility regulation in the UK has relied heavily on a building block approach to setting the price path under a RPI-X regime. At its most simplistic, a building blocks approach involves splitting a regulated company's revenue requirements into different categories and assessing them separately. The highest level of assessment is more often than not conducted on the following three categories: return on their regulatory asset base; operating expenditure (opex); and capital expenditure (capex), including depreciation allowance. Other building blocks (e.g. tax, non-regulated revenue) can be assessed depending on the needs of the regulator, or the requirements of the regime. The RPI-X form of control used by many regulators contains numerous incentive mechanisms designed to incentivise quality and innovation, as well as equalising incentives across opex and capex (where these expenditure types are assessed separately).

Regulators usually define companies' revenue requirements through the three broad categories, but often require companies to report opex and capex in greater granularity, e.g. expenditure on IT; fault repairs; maintenance; replacement capital expenditure; capital expenditure related to load growth; etc. In addition to reviewing expenditure at the more granular level, regulators have required companies to provide 'splits' of opex and capex for various stages of the value chain, e.g. large sewage treatment plants. The extent to which companies are required to provide cost splits across their different business units varies across regulators. Figure 4.1 below illustrates how the cost 'blocks' can be built up across a regulated business. Analysis and/ or performance measurement is theoretically possible for each block.

Figure 4.1 – Building blocks



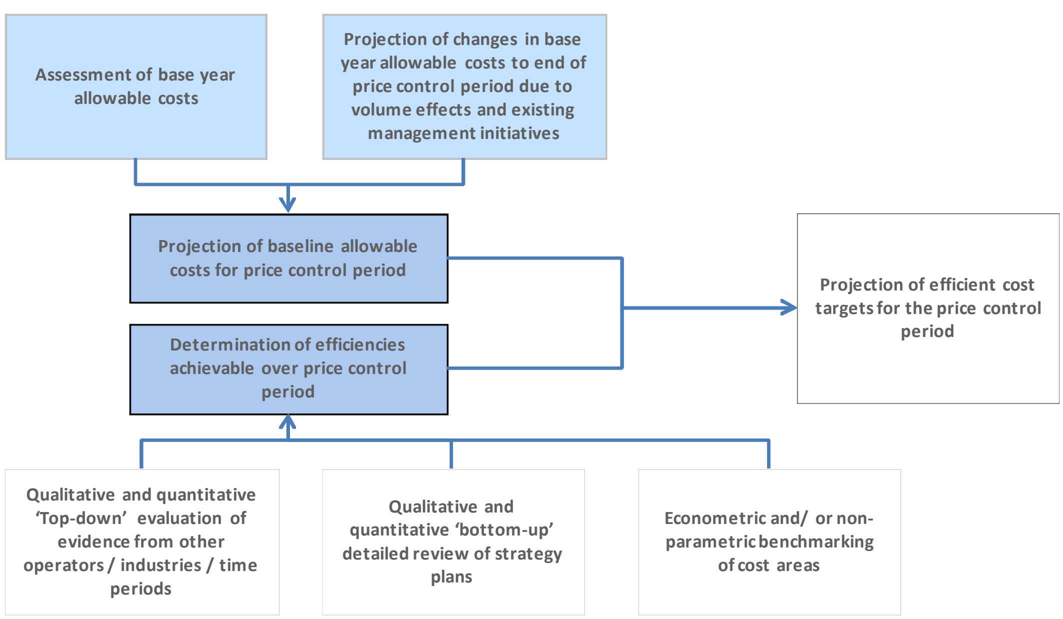
A building blocks approach assists in setting a price path through: (i) determining allowable costs for a base year; and (ii) estimating the scope for efficiency improvements over the price control period. These efficiency adjustments can be applied at the start of the price control period (sometimes referred to as a P_0 adjustment) or spread across the price control period. When the regulator has determined the revenue requirements for each separate category of costs, they aggregate these up in order to determine the efficient cost path for each company over the period of the price control. The overriding assumption of the modelling, that the cost assessment is conducted on comparable companies, infers that the same cost interactions would exist across the companies. This assumption allows this approach to be adopted no matter what economies of scale, scope or integration exist. As long as companies have similar cost structures then it can be assumed that cost interactions are implicitly included. Figure 4.2 below illustrates how an efficient cost path can be set.

Historically regulators have focused on using actual data for benchmarking purposes and relied on forecasts in order to help project base year costs forward. However, Ofgem has proposed that it will use forward looking data in its benchmarking for the next gas distribution and transmission price control reviews.

4.2.2. Approach to comparative analysis

The lower half of Figure 4.2 illustrates some of the different approaches that can be used by regulators in determining the scope for achievable efficiency improvements (both catch-up and technical change). These approaches include qualitative, relatively simplistic quantitative assessments and more sophisticated benchmarking approaches. The more simplistic quantitative assessments include: (a) reviewing simple financial ratios; and (b) unit cost comparative analysis. We focus less on these approaches given their relative simplicity, (generally) lower data requirements and reduced model specification, i.e. these approaches can be conducted with few(er) observations and, if the correct ratios are chosen, on sub-company data.

Figure 4.2: Setting an efficient cost path



Regulators have used a number of different methodologies in benchmarking relative efficiency and technical change, the methods used include COLS, SFA, DEA, and Indices. A discussion on the advantages and disadvantages of these approaches was set out in Section 3.3.

While regulators make use of statistical tools for benchmarking, the analysis has generally been done:

- with less sophisticated methodological approaches than those more recently employed by academics;
- with various cost drivers;
- at different levels of aggregation; and
- with cross-sectional analysis or shorter panels than those used in more recent academic studies (as discussed in Annex A).

A regulator's decisions about the best approach to benchmarking involves implicit cost benefit analysis regarding the amount of resources (both in terms of the regulator and the regulated companies) to devote and the approaches to be adopted. Consideration is particularly given to achieving a balance between accuracy and discretion (e.g. allowing for outperformance). Regulators are also – and understandably – cautious about relying on sophisticated statistical methods that are difficult to explain to non-experts and which depend for their reliability on strong assumptions.

In general, while it may be possible to justify committing a relatively large amount of resources on the basis of possible efficiency savings, a more sophisticated set of considerations are relevant, including:

- the robustness of the results that are likely to arise from each approach;
- the number of observations, and the accuracy and stability of available (or requested) data;
- the transparency of the approach used and its reproducibility; and
- differences in the operating environment (heterogeneity) of the businesses, including historic network/ asset setup.

Where regulators have used econometric, DEA or index models, they have generally adopted a number of approaches to mitigate potential issues in the model (e.g. misspecification, measurement error, etc) on the modelling results and how these impact on the regulated companies. These approaches have included:

- setting an efficiency target that is less challenging than indicated by the modelling; and
- taking a more discretionary approach to using the modelling outputs to set efficiency targets, i.e. using them to inform the targets rather than deterministically setting them.

In addition, we consider that Ofwat should treat evidence from comparative analysis carried out at lower levels of costs with care as disaggregation of the data can highlight comparability issues between the benchmark companies/ decision making units. Section 3 provided a discussion on some of these issues.

In the following sections we present examples of regulators' use of comparative analysis, and more specifically their use of sub-company and panel data in the UK and internationally.

4.3. Ofwat's PR09 approach

At PR09 Ofwat set price limits for 10 WaSCs and 12 WoCs, however Cholderton Water was excluded from the comparative analysis due to its small size. Ofwat relied on building block analysis to determine the revenue allowance for the WaSCs and WoCs. Ofwat analysed opex and capex costs separately – as explained below.

Relative efficiency methodology – operating costs

Ofwat's most extensive use of econometric techniques in analysing efficiency covers the operating expenditure of each of the companies. At PR09 Ofwat relied on COLS and unit cost models applied to cross-section data for 2008-09 to determine an overall efficiency level and benchmark. For the water service Ofwat employed four different functional models (based on specific operating activities of the companies) at a sub-company level (where a sub-company is considered to be a vertical segment of the companies' value chain) to assess relative operating cost efficiencies across the water companies and the water side of the WaSCs. The models were for:

- water distribution;
- water resources and treatment;
- water power; and
- water business activities.

Ofwat combined the results of these models to give an overall water service operating expenditure efficiency band.

On the sewerage side, Ofwat employed two **sub-activity based** COLS operating expenditure models (where sub-activity refers to an activity at a sub-company level, for example, large treatment plants) and three unit cost models to determine overall efficiency scores. Again these models used **one year of cross-section data**. The models covered were:

- sewerage network (including power);
- large treatment works;
- small treatment works unit costs;
- sludge treatment and disposal unit costs; and
- sewerage business activities unit costs.

As with water services the sewerage models Ofwat combined the results from the models to give an overall sewerage service operating expenditure efficiency band. Ofwat made adjustments to ensure consistency and comparability between the companies. These included an adjustment for capex-opex interaction and for pensions costs. Ofwat combined these adjustments with the modelling results in order to determine the overall efficiency levels for WaSCs and WoCs.

While Ofwat relied on statistical and financial methods in its benchmarking it employed a number of techniques that compensated for possible measurement or misspecification (e.g. explanatory variables not included) errors. These reduced the emphasis on using the resulting overall efficiency scores in a deterministic way.

The main allowances Ofwat made in the application of its modelling in PR09 were to: (i) reduce the residuals to take into account the possible errors in the data and the statistical process; (ii) determine the benchmark company taking into account other factors about the company; and (iii) use bands for companies' efficiency based on their distance from the benchmark. If companies were close to the upper border of the band then Ofwat promoted them to the next band - which resulted in a lower efficiency target.

In addition to relative efficiency analysis (catch-up), Ofwat engaged Reckon LLP to analyse the scope for frontier shift or a 'continuing improvement factor'.[33] Ofwat set separate continuing improvement factors for total capex and opex.

Relative efficiency methodology – capital expenditure

For PR09 Ofwat did not use any econometric cost modelling for capex (unlike previous price controls). Instead Ofwat used the cost base comparative tool, as used in previous price controls, and the capex incentive scheme (CIS).

The cost base comparative tool was used to assess relative efficiency in companies' procurement and delivery of capital projects. The cost base comparative tool works by comparing company estimates of capital works unit costs for a representative range of standardised capital projects.[34] This analysis was used to inform a base line level of capex which then formed the starting level of the CIS.

The CIS is not in itself a relative efficiency assessment tool, rather it is a menu regulation mechanism which was used by Ofwat to set expenditure assumptions and associated rewards for outperformance, for capital maintenance and enhancement expenditure. Menu regulation works by offering companies a trade-off between the allowed level of capex and the benefits/penalties associated with the target and under-/over-performance.

4.4. Other regulators' approaches

As noted at the start of this section, the building blocks approach is a common approach used by regulators in the UK. Below, we have set out a number of examples of recent price controls where the regulators have used one or both **sub-company** or **panel data**. We also provide an example of a non-building blocks approach to setting relative efficiency targets that was used by the New Zealand (NZ) Commerce Commission for its 2004-09 electricity distribution businesses price control review.

[33] Ofwat engaged Reckon LLP as part of PR09 to produce a report investigating the scope for efficiency studies, see Reckon (2008).

[34] Ofwat (2009).

Ofgem

Ofgem's approach during its most recent price control review (DPCR5) relied heavily on **panel data** for **sub-activities** for setting both opex and capex allowances.

To inform its network operating and indirect cost allowances for DPCR5, Ofgem applied corrected ordinary least squares (COLS) based on a panel dataset. The use of a panel dataset allowed for 56 observations to be included in the modelling (14 distribution network operators [DNOs] with four time periods each [2006-2009]), rather than a cross-section of 14.

In the first instance, Ofgem undertook OLS to determine the weights for a composite scale variable (CSV) to use as the driver to regress the costs against. Given the group ownership existing across a number of licences, Ofgem also conducted regressions on indirect (e.g. overheads) costs where costs were grouped together by DNO ownership. There were eight separate 'blocks' of regressions with the cost base, driver, driver weights and outliers varying for each. The drivers used by Ofgem varied from MEA value to the number of spans where tree cutting occurred. While a CSV was created for most cost items, some costs/ groups only had one driver.

These variations resulted in over 40 separate efficiency scores being produced for each DNO. Ofgem weighted these scores together to form an overall efficiency score, a network operating costs (NOC) efficiency score and an indirect costs efficiency score. The difference between the efficiency scores and the benchmark, based on either the upper third percentile or upper quartile (for NOC and indirects respectively), was used to determine an efficient baseline for each DNO. We consider that the approach taken by Ofgem, i.e. using efficiency assessment at group, company and sub-activity levels, and weighting the resulting scores together (but placing more weight on the high aggregation modelling), in effect took account of the cost interactions occurring at the more granular level. In other words, Ofgem put less weight on sub-activities as assessment at this level does not allow for the full scope for efficiency improvements.

For capex, Ofgem used panel data to analyse unit costs and volumes to determine allowances for load related capex, replacement capex (depreciation), and connection expenditure. There were other costs areas which Ofgem assessed, but we have focused on the main areas where sub-company or panel data was used. This was then incorporated into the IQI, the Ofgem version of a menu.

A recent change in Ofgem's approach to efficiency assessment has been brought about with Ofgem's move away from a RPI-X based regime to "Regulation = Innovation + Incentive + Outputs" (RIIO). The 'toolkit' methodologies set out for RIIO focus on totex benchmarking first and then if it is deemed necessary Ofgem will drill down on a step-by-step basis into further detail (i.e. from totex benchmarking to project by project review), this staged approach is shown in Table 4.1 below.

In addition to moving toward totex benchmarking, Ofgem has announced that as part of the next transmission price control review (T1) and gas distribution price control review (GD1) it will conduct benchmarking on *forecast* costs.[35] As discussed in Section 2.3, it is not yet clear the extent to which Ofgem intends to use forecast costs in its benchmarking. A report

[35] Ofgem (2011b).

commissioned by Ofgem, *RPI-X@20: The future role of benchmarking in regulatory reviews*,[36] notes that benchmarking future plans can:

- minimise distortion of opex-capex trade-off;
- give a high level assessment of value for money for customers; and,
- reduce the regulatory risk of stranding assets through the assessment of future costs against future output.

However, the report also notes risks associated with this approach including: an incentive to inflate plans; and uncertainty over future activity. Given the onus that RIIO places on long-term value for money and business plans, Ofgem does not intend to use benchmarking in a mechanistic way, rather it will be used to inform Ofgem's assessment of companies cost forecasts.

Table 4.1: RIIO Assessment toolkit

Source: Ofgem, *Handbook for implementing the RIIO model,* October 2010

Ofgem intends to use a COLS methodology for the transmission price control review rather than SFA given that "it provides more reliable estimates than SFA when applied to small sample sizes."[37] Ofgem stated that it preferred COLS to DEA as the underlying statistical properties of the method allows for tests on the reliability of the estimates. Ofgem intends to use DEA to cross check the results of the COLS modelling.

ORR

ORR's efficiency benchmarking approach for PR08 considered total maintenance and renewal expenditure. SFA was the main technique used, but checks were performed with COLS and DEA. ORR used an 11 year **panel dataset** of international rail network operators (12 other European operators) with which to benchmark Network Rail against. As well as undertaking a 'top-down' econometric analysis of Network Rail's costs, ORR also carried out a 'bottom-up'

[36] Frontier (2010).
[37] Ofgem (2011a).

engineering based review which it relied on to validate and better understand the efficiency targets produced via the econometric analysis.

The overall efficiency saving determined and imposed during the price control by ORR was an important part of setting the access prices that Network Rail charges train operating companies. The efficiency assessment has particular impact on determining the variable usage charge that is part of the overall access charge.

Postcomm

Postcomm is responsible for regulating one company, Royal Mail. As part of the 2005 price review LECG (an economic consultancy) were engaged by Postcomm to review the scope for efficiency savings in Royal Mail. LECG undertook comprehensive top-down and bottom-up reviews of Royal Mail costs. It carried out internal benchmarking across Royal Mails' mail centres (70) and delivery offices (1383) using DEA analysis and econometric frontier analysis. The analysis was conducted on a single year of data (2003/04).

LECG also undertook four different approaches to the top-down assessment. These approaches included comparing TFP ratios to other international mail operators and other regulated sectors in the UK.[38] LECG used previously determined real unit operating costs (RUOCs) for the comparators, therefore the period covered by the TFP indexes varied. There was also significant variation in the growth rates recorded for international comparators ranging from -7% to 7.4% (compound annual growth rate, adjusted for volume differentials). The degree of variation across the international comparators indicated that this measure was not reliable.

New Zealand Commerce Commission

As part of its second price control review (2004-09) for electricity distribution business the New Zealand (NZ) Commerce Commission engaged Meyrick (an economic consultancy) to undertake a review of the scope for industry productivity improvements, relative efficiency improvements and profitability. The Meyrick study relied on a panel data set covering 29 distribution businesses over six years. The Meyrick study used a parametric index number approach that allowed for the determination of both a firm's relative efficiency level and its growth over time. Based on these results a business was placed into one of three bands which either required no change in efficiency, relatively faster efficiency gains or relatively slower efficiency gains. While this technique provided insight into the different efficiency levels of the distribution business, the businesses cited issues with the robustness of the results and the predictability of the modelling.[39] The regulator noted the concerns raised, but on balance considered that the proposed approach was appropriate for setting the price paths.

Interestingly, the Act under which the 2004 price control was set has since been amended and a 'default price path' is now set for the electricity distribution businesses. The amendments stipulate that comparative analysis cannot be used in setting the default price path. The

---

[38] The TFP ratio was based on a real unit operating cost (RUOC) ratio, which is equal to firm TFP growth *less* economy-wide growth *less* an input price adjustment.

[39] A cost based function was used to determine weights for the outputs. Given the use of panel data this increased the scope for companies' efficiency rankings to retrospectively change compared to estimates determined using a different length panel data set.

electricity distribution operators concerns about the transparency, robustness and appropriateness of the comparative benchmarking approach were again raised as part of the consultation process reviewing the Act and are likely to be one of the reasons for the change.

Table 4.1: Regulatory examples – Relative operating cost efficiency assessment

Sample	**Ofwat – PR09**	**Ofgem – DPCR5**	**ORR – PR08**	**Postcomm – 2003**	**NZ Commerce Commission – Electricity distribution price control 2004-09**
Methodology	COLS and unit cost models.	COLS.	SFA analysis. COLS and DEA used as a cross check	DEA, econometric frontier analysis and SFA used. TFP ratio for top-down analysis.	Multilateral TFP analysis(Index)
Sub-company – vertical segment	Yes. Models used to assess different segments of the water value-chain and the sewerage value-chain.	No. Network companies are legally separated from other parts of the value-chain. However, multiple licence ownership by some companies.	No. Single regulated company with separate assessment of vertical companies.	No. Single regulated company with separate assessment of vertical companies.	No. Network companies are legally separated from other parts of the value-chain. Analysis was done on aggregate expenditure.
Sub-company – cost type	Yes. Assessment conducted on activities/ cost areas e.g. large treatment works, power costs, etc.	Yes (sub-activity). Different cost types are modelled separately i.e. tree cutting, IT, property etc.	No.	Yes (sub-activity). Benchmarking using DEA was carried out on the Royal Mail's deliver centres and mail offices.	No.
Panel	No. Single cross-section used.	Yes. Four years of data.	Yes. 13 years of data from 14 European rail infrastructure managers.	No (for internal benchmarking). RUOC growth rates across a number of years from different countries and sectors were used for top-down benchmarking, but were not calculated in the same model.	Yes. Six years of data (Companies were asked to back fill data and consistency issues were identified).

5. OUR APPROACH

Key themes and lessons

Uncertainty about the future industry structure means it is appropriate to develop a flexible assessment approach. This is built around a ***decision tree***.

The two key questions that need to be considered are:

1. At what level of the value chain will the price control(s) be set? And
2. What is (are) the chosen, corresponding performance measurement requirements?

With respect to the latter, the increasing focus on competition in the sector means that performance measurement to facilitate competition, either through increased transparency of cost data or by ensuring efficient access prices, is more appropriate.

Consequently this means that there are three possible uses of performance measurement:

1. determining efficient costs for price determinations;
2. facilitating and monitoring competition; and
3. providing a consistency check on costs within the sector.

In this section we present our approach to determining the feasibility of using sub-company and/or panel data with different models under various regulatory structures. Our approach is to create a decision tree based around key questions and selection factors. This provides a tool that can be adapted to different industry/regulatory structures as the industry develops. Our approach relies on the concepts and issues set out in Section 3 above (and Annex A).

5.1. Decision tree

The industry/regulatory structure is a key determinant of overall requirements for performance measurement and as such is the starting point for any evaluation. Given the potential scope for changes to the existing industry structure and hence the level along the value chain at which Ofwat applies cost efficiency modelling, we have developed a simple decision tree for determining the feasibility of different performance measurement options. Consequently the choice of industry structure is the first step on the decision tree.

The shape of the decision tree for a given industry structure is driven by two key questions:

- At what level of the value chain will the price control(s) be set?[40] And
- What is (are) the performance measurement requirement(s)?

Based on the answers to the preceding questions, a second 'stage' of the decision tree determines the feasibility of performance measurement at this level.

[40] This is concerned with both the legal price control and any indicative controls that may be set.

5.2. At what level of the value chain will the price control be set?

The level at which price caps are set determines the minimum aggregated cost level at which performance measurement should be conducted. It is important to note that while the price cap level may be set at, for example, the level of the regulated company, this does not preclude performance measurement being undertaken on sub-company or sub-activity data. The price cap reflects the level at which the performance measurement should be undertaken, but further information may be gathered through performance measurement at a more disaggregated level. In contrast, conducting performance assessment at a lower level of aggregation, generally necessitates additional cost assessment at a higher level to ensure consistency of the estimated costs.

We note that Ofwat is considering having more than one price cap when it sets price limits in the future.[41]

5.3. What is (are) the performance measurement requirement(s)?

The next step on the decision tree is to consider the purpose of undertaking performance measurements. We believe that the requirements for the performance measure fall into three broad categories, which are not mutually exclusive.

5.3.1. Efficient cost determination

Efficient cost determination refers to the use of the performance measure to assist Ofwat in setting efficiency targets for the companies over the price control period. Within this use, the performance measures assess the following:

- the potential for efficiency catch-up by lagging companies;
- the potential for continuing efficiency change across the industry as a whole (TFP growth at the frontier); and/ or
- appropriate access prices for a monopoly or essential facility.[42]

Efficient cost determination is the most common use of a performance measure and it is theoretically possible to use some form of measurement at any level of the value chain. Although, as explained in Section 3, the fact that it can be applied does not mean that a meaningful or unbiased estimate of efficiency will be found from sub-company estimation. If economies of scale or scope exist, then a biased estimate may be found and Ofwat may also need to undertake the measurement at a different level of aggregation *and*/or consider some form of adjustment to the modelling or results.

[41] Ofwat (2011a).

[42] Water pipes are typically natural monopoly facilities whereas treatment works and similar are often essential facilities but not natural monopolies.

5.3.2. Facilitating efficient competition

The use of the performance measure for facilitating efficient competition is dependent on the industry structure and the objectives of the regulation. At the given level where it is applied, the performance measure could be used to provide further information on:

- the regulated incumbent's costs in potentially contestable markets; and/ or
- the appropriate access prices for a monopoly or essential facility.

Given the nature of the industry, the scope for using the performance measure will be restricted to areas where contestability has been identified and may be viable over time (not necessarily in the short-term).

For example, as suggested in Ofwat's January 2011 assessment of hypothetical upstream water markets,[43] efficient competition will require access pricing at the sub-company level, and location specific information on an incumbent's marginal costs in potentially contestable markets. Sub-company modelling, would in principle provide useful information to meet both of these needs.

5.3.3. Consistency checks

A performance measure can be used for two forms of consistency checks:

- *Cost interaction.* To ensure that cost interactions and cost allocations between activities of an integrated firm which are assessed at a lower level of vertical or horizontal aggregation have not resulted in a biased overall cost assessment (discussed in Section 3 and Annex A).
- *Performance assessment (continuity).* A consistency check between past performance assessment measures and new forms of performance assessment under development.

Simply put, the first ensures that the sum of the parts equals the whole; while the second checks for consistency in results across different models. The first check may be done on a level above that at which the price cap is set. This would be done where the price cap is set at a level below company ownership, or potentially where there is group ownership of multiple companies within the regulated segment.

In an ideal world the second objective would be achieved by running the old and new models concurrently (as Ofwat has done in the past when changes occur) to check for robust consistent results. It may not be possible, however, to do this and consequently an approach based around running several alternative approaches, or some higher level checks, could be considered to check the robustness of the new model. This issue is discussed in more detail in Section 9.

5.3.4. Possible performance measurement requirements for PR14

Table 5.1 below illustrates the possible requirements for performance measures for PR14 if efficiency modelling were to be conducted based on data provided in line with the current accounting separation guidelines. As noted above, performance measurement can feasibly be

[43] Ofwat (2011b).

used at any level of the value chain for efficient cost determination. The use of performance measurement for facilitating competition is, however, less useful at an aggregated level, i.e. it is unlikely competition could be introduced, by PR14, at the current regulated business level or the potential regulated business level. A 'higher level' performance measure could be used to help set efficient access prices, but the appropriateness of this is dependent on the level of heterogeneity at the access level.

Performance measurement can be used as a consistency check at any cost aggregation level above sub-activity. For PR14, the intertemporal performance measurement check can be done at the current regulated business level. It is also feasible that Ofwat's current opex sub-company models could be used as a partial consistency check, depending on the similarities in the structure of the cost data between PR14 and PR09.

Table 5.1: Example of the possible performance measurement requirements at PR14

Performance measurement	**Current regulated business (i.e. WaSC or WoC)**	**Potential regulated business (e.g. Water excluding retail)**	**Vertical segment (e.g. Water treatment)**	**Sub-activity (e.g. large sewage treatment plant)**
Efficient cost determination	✓	✓	✓	✓
Facilitating competition	×	×	✓	✓
Consistency - cost interaction	✓	✓	✓	×
Consistency – performance assessment *(continuity)*	✓	×	×	×

5.4. Data availability

The first consideration for the option assessment is to determine the extent to which data is available to undertake the performance measurement.

Data collection at levels of disaggregation below those traditionally reported in regulatory accounts (or in company accounts), are likely to suffer in the near future from low accuracy and stability. (Anecdotal evidence would suggest that at least two years need to elapse before the data stabilises at lower levels of disaggregation.) Given the changes to data collection imposed by Ofwat's new accounting separation guidelines in 2009/10 it is likely that only data at the highest level(s) of aggregation indicated in Table 5.1 (i.e. current regulated business) will be available for significant amounts of panel analysis at PR 14. So, when determining whether appropriate data is available for performance measurement the length of the dataset needs to be considered against the reliability of early observations. Ofwat will need to decide whether it is confident enough about the accuracy and stability of the data to use it in its cost modelling.

Given the discussion above the key questions at this step in the decision tree are:

- Are there a sufficient number of observations available? And

- Are there a sufficient number of comparators available?

As discussed in Sections 3.4, 3.5, and Annex A, modelling at a sub-company or sub-activity level may require a number of additional variables to control for heterogeneity, scale effects etc. As the number of comparators at the sub-company level are limited (and may reduce through future mergers although this is not necessarily the case e.g. the number of treatment plants), this would suggest that panel or pooled data would be required to provided sufficient observations to produce robust estimates. While there are likely to be more 'comparators' at the sub-activity level, pooled or panel data may be a requirement as well if insufficient numbers are available.

As we have posited in Section 3.4 and Annex A.6, if shorter datasets (e.g. less than four years) are only available then applying panel models on a pooled basis may be applicable if tests for the presence of technical change prove to be statistically insignificant. However, if a longer database (of consistent comparators) is available then we would suggest that any panel cost modelling be conducted with controls for technical changes. If technical change is present, but only a short data set is available, then year dummy variables may be sufficient to control for technical change. However, this would require the assumption that the time specific effect is individually invariant (i.e. all companies experience the same year-on-year change).[44]

5.5. Performance measurement methodology

The choice of performance measurement is driven by the requirements and objectives of Ofwat. This is an important decision, and we have discussed the different approaches available to Ofwat in Section 3.4, however we do not propose an in depth review of the various different methodologies at each level. Rather we need to consider, at a high level, the appropriateness of the specific types of methodologies – econometric, DEA and non-parametric index – at the given level of the value chain.

Table 5.2 below indicates the likely feasibility of different methodologies at different levels of the value chain. This assessment is based on the assumption that there is suitable data available (i.e. the previous step of the decision tree has been passed). As can be seen from the table, it is our view that econometric methods and DEA are feasible at all levels of the value chain – at least on a cross-sectional basis and possibly with pooled, non-panel data. However, it needs to be recognised that the potential for biased estimation increases substantially as one moves to greater levels of disaggregation. This is primarily because of the various issues discussed earlier regarding the impact of potential vertical and horizontal linkages. We do not consider that non-parametric index number approaches are likely to be feasible at lower levels of the value chain because of their requirements for measuring capex or total cost, their relative inflexibility in dealing with different operating characteristics, and their requirement for output price proxies.

[44] Please see Section A.7.1 for further discussion on issues around technical change in a pooled model.

Table 5.2: Feasible methodologies at different levels of the value-chain

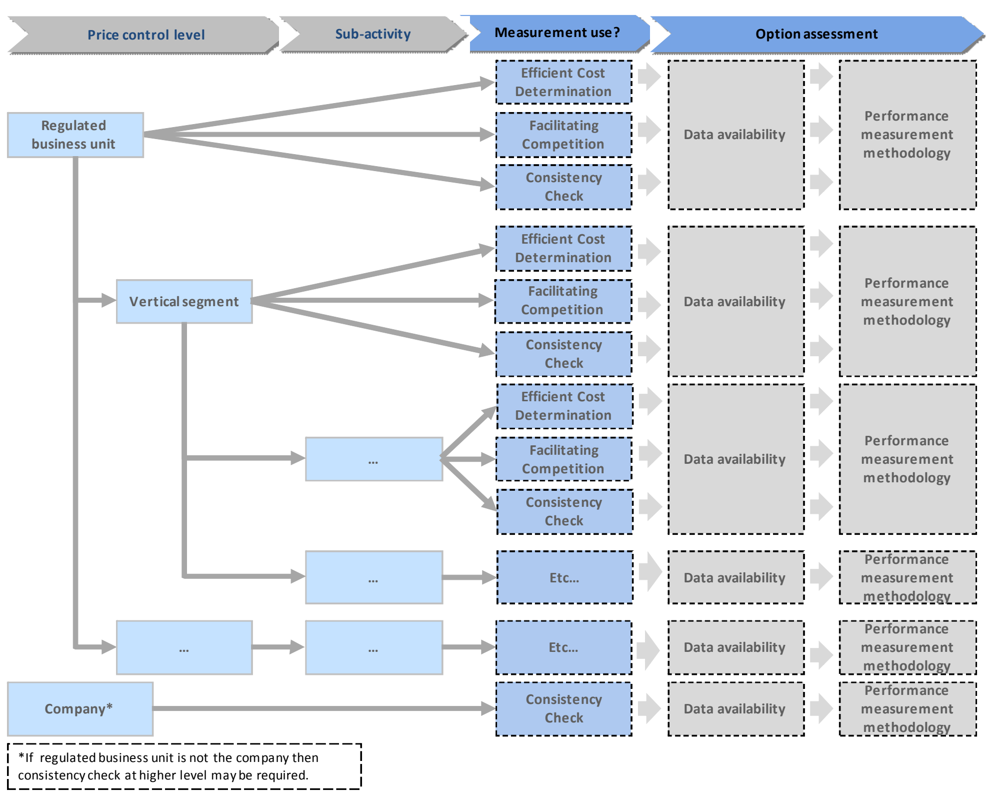
Performance methodology	Current Regulated Business	Potential Regulated Business	Vertical Segment	Sub Activity Level
Econometrics (COLS, SFA)	✓	✓	✓	✓
DEA	✓	✓	✓	✓
Non-parametric index (Tournqvist, Fisher)	✓	✓	×	×

5.6. Summary

This section sets out our proposed decision making steps and options assessment for determining the appropriateness of using panel, sub-company and/or performance measurement at different levels of the value chain. Figure 5.1 below illustrates the decision making process (i.e. Decision tree) for a given industry structure. The first three stages ('price control level' to 'measurement use?') of the decision tree determine the requirement for sub-company data. During the option assessment stage the use of panel data and feasible methodologies are considered.

In the next section we demonstrate how this decision process could be applied. We have used an example of the data that companies report to Ofwat in line with the current accounting separation guidelines.

Figure 5.1: Decision tree



6. POTENTIAL APPLICATION AT PR14

Key themes and lessons

The companies currently report their costs under nine business units using the accounting separation guidelines. Based on existing Ofwat statements,[45] we understand that at PR14 it may set separate wholesale and retail price controls but with possible sub-caps for some of the business units within the wholesale control. Consequently:

- Performance measurement for price setting is likely to be needed across all the industry segments although this does not necessarily mean that a price cap will be set for each business unit;
- Performance measurement at sub-company and sub-activity levels may be needed in a small number of areas to facilitate competition, this is more likely to be important for sludge treatment and disposal, water resources and possibly competitive retail;
- Depending on policy around the speed of the growth of more general competition there may be a need for Ofwat to undertake performance measurement at the sub-company level for distribution activities to be able to set efficient access prices, although this may be something that occurs beyond the PR14 determination;
- It may be necessary for Ofwat to undertake consistency checks across a range of levels, including at the company level, given potential issues around cost allocation, cost interactions, etc; and
- We believe that most forms of performance measurement are possible at PR14, however, we consider that non-parametric index number approaches are less desirable, than econometric or DEA approaches, at sub-company or sub-activity level as there is less scope in these models to control for heterogeneity and the requirement of output price indices. (Note that we do not discuss further whether separate opex and capex or totex benchmarking should be undertaken as this is outside the remit of this project.)

The existing price control was set at the level of the entire business for WoCs and WaSCs (separate indicative price controls were set for water and sewerage services), with efficiency analysis being conducted separately for water and sewerage. Ofwat used models for this efficiency analysis which varied from being conducted on vertical segments within each company type (e.g. water distribution) to sub-activity (e.g. large sewage treatment plants). Ofwat has introduced accounting separation guidelines that require regulated companies to provide accounting information split between nine different business units – four water, four sewerage and one retail.

Modelling using accounting separation data is a logical extension of the modelling approach behind Ofwat's PR09 benchmarking approaches as well as its sub-company data approach for performance measurement, e.g. breaking the whole into its parts allowing better understanding of the cost drivers. Moreover, if the vertical lines between units of assessment are drawn "appropriately" accounting separation allows assessment of "parts" which may be economically valid "wholes". However, as explained in Section 3, if the levels chosen are not appropriate then

[45] Ofwat (2011a).

biased estimates may be found and Ofwat would need to undertake the consistency checks discussed earlier to determine if controls were needed for cost allocation issues.

Accounting separation may provide sufficient data to allow Ofwat to implement sub-company price caps in 2014 or beyond. However, without legal separation, estimating the costs of integrated operations may suffer from cost allocation issues and be subject to potential regulatory gaming. Even with legal separation, if costs are non-separable and/or if substantial vertical integration (dis)economies exist between all or some accounting separation parts, failing to properly control for these cost interactions may well result in potentially biased cost assessments. This may still be the case even if efficient cost transfers were in place, as economies of scope and scale may still exits. However, if all companies have the same organisational structure then although the performance measure may be biased, it is biased for everyone and it is not an issue when comparing companies or assessing their costs. If some companies have a different organisational structure then it is possible that the bias in the cost assessment will be uneven.

Ofgem does not appear to carry out consistency checks across the vertical segments in the energy value chain. This may be because Ofgem is satisfied that cost interactions in the value chain are not creating biased costs assessment (for example, the electricity distribution business may be close to the optimal size [or size range]).

In this section we place the greater focus on how the decision tree applies to water supply, but with some consideration of sewerage and retail in the later sub-sections. This emphasis is because the same types of issues arise no matter which of the segments of the industry are being evaluated. Under the accounting separation guidelines, companies' retail business units are required to be reported separately. Retail business units are considered in Section 6.3.

For the purpose of illustrating the approach set out in Section 5 in the analysis below we have assumed certain price cap levels and activity breakdowns for PR14. These assumptions are based on plausible, but not necessarily probable, outcomes given Ofwat's accounting separation guidelines. In addition, we also make the following two assumptions:

- no further competition would have developed by PR14, however, if it were to develop then the access pricing issues discussed in Section 7 would need to be considered; and
- the allocation of capital and the privatisation discount between business units will have been satisfactorily dealt with. Our analysis of data availability and performance measurement feasibility does not take in to account the capital allocations and privatisation discounts.[46]

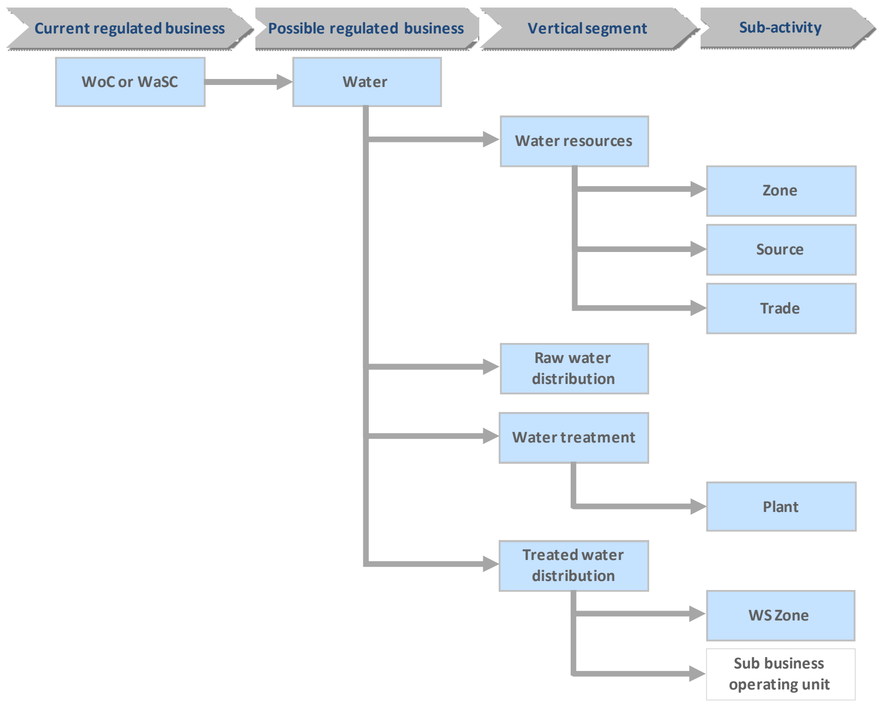
[46] The privatisation discount is the discount provided on the asset base when the companies were privatised.

6.1. Water

6.1.1. Possible structure

Figure 6.1 provides an illustration of the possible structure of the water sector under Ofwat's accounting separation guidelines.[47] The vertical segments are based on Ofwat's business units, while the sub-activity level illustrates plausible (not necessarily probable) smaller decision making units (DMUs). Note, the *water* structure excludes *retail*, which is discussed in Section 6.3.

Figure 6.1: Possible structure of 'water'



While the sub-activity splits are for illustrative purposes, it is plausible that the vertical segments could be broken down in to further DMUs. Sub-activity DMUs in Figure 6.1 refers to:

- *Zone*. The resource zone in which all water resources, including external transfers, can be shared.[48]
- *Source*. The DMU in control of abstracting from a source (e.g. aquifers, lakes, reservoirs, rivers and third parties).
- *Trade*. Covers the DMUs involved in trading bulk supplies of treated or untreated water.
- *Plant*. The 'plant' involved in the treatment of raw water (or partially treated raw water).
- *WS Zone*. Water supply zone.
- *Sub business operating unit*. A functional business unit with the ability to make decision regarding its own activities, such as distinct water supply units within a regulated firm.

[47] Summaries of the business units defined by Ofwat are set out in Appendix 1 of the accounting separation guidelines, see http://www.ofwat.gov.uk/competition/rrq_jr09-10_acountingsepappen1.pdf.

[48] Ofwat's definition of resource zone is set out in the glossary for the accounting separation guidelines, see http://www.ofwat.gov.uk/regulating/reporting/pap_tec_lrmcglossary.pdf.

Of course, were further competition to be introduced over time, some sub-activities may become more dependent on the form and level of access prices. This issue is considered in Section 7.

6.1.2. Evaluation

As noted in Section 5, the requirements for performance measurement are determined by the industry structure, requirements for the regulatory regime and Ofwat's objectives.

We have applied the decision tree to what we consider to be plausible outcomes under Ofwat's existing accounting separation guidelines at PR14. In order to restrict the number of independent assessments we have grouped the different DMUs at similar levels across the water business. The summary results for these are set out in Tables 6.1 and 6.2.

Water

We have assumed that this is the level that Ofwat will set the price controls. At this level, Ofwat will need to use performance measures for efficient cost determination and as a consistency check (for cost allocation issues). Moreover, for water operations contained within a WaSC, Ofwat should carry out higher level consistency checks to properly control for cost interactions between water and sewerage activities. Ofwat will not need to use the performance measure at this level to facilitate competition as this is infeasible by PR14 at this level of the value chain.

For PR14, we consider that there is likely to be a stable dataset with approximately three to four years of data available as: (i) the accounting separation guidelines have been in place since the 2009/10 reporting year; and (ii) as this is a high level of the value chain (e.g. high level of cost aggregation), cost allocation issues will be relatively minor. A model based on panel data is possible with a dataset of this length and breadth, however Ofwat may also consider the use of a pooled data specification, if it determines that technical change is not significant, rather than panel (i.e. pool all observations together and run through a standard OLS model). If Ofwat determines that there is sufficient data, then given the benefits outlined in Section 3.5, e.g. the ability to control for cross sectional and intertemporal heterogeneity, we would recommend a panel based model. At this level there may be potential for Ofwat to conduct assessments using data from the pre-accounting separation guidelines period, thereby allowing it the use of a longer panel. However, Ofwat should carefully consider and test the data, and controls for structural breaks may be necessary to ensure that structural breaks in the data did not influence the estimated costs.

Table 6.1 below summarises the key factors that Ofwat needs to consider and potential implications and solutions. Ofwat may need to introduce a number of variables to control for various factors (e.g. heterogeneity) at this level, if this is the case then a sufficient number of observations will be needed to provide robust modelling estimates. Panel data may be required to provide sufficient degrees of freedom.

At this level of the value chain any performance measurement methodology is feasible, including non-parametric index numbers. While non-parametric index numbers provide less scope to control for different operating characteristics, if Ofwat determined this was not a significant issue at this level then it would be suitable to use them.[49] However, the choice of methodology is dependent on Ofwat's objectives. Requesting data on variables potentially required to control for heterogeneity, scale, etc, would create additional regulatory burdens on companies and Ofwat. A number of variables that might potentially be needed are, however, currently collected as part of the June return process. Therefore, we believe this issue is more a matter of being careful to retain necessary control variables, if Ofwat were to revise the required content of the June return to reduce regulatory burden.

Table 6.1: Key considerations for performance measurement for water (excluding retail) at PR14

Issue	**Relevance to this part of the value chain**	**Implication and solutions**
Underlying economic theory		
Input prices	Input price differentials across *water* businesses are likely to be mainly driven by geographical differences (e.g. one region's labour costs are relatively lower than in another region). The possibility of a longer dataset at this level allows the possibility of capturing input substitution over time as prices change. Intertemporal price changes will also need to be considered.	At this level Ofwat could deal with relative input price differentials through regional price adjustments (e.g. ONS regional wage data). Ofwat could also include controls in the modelling to allow for input substitution. Alternatively, it could make adjustments to the input data or post-model (i.e. discretionary adjustment to the allowance) to take account of these differences. Both approaches require information on price changes across the regions (and over time) which, in some cases, may be difficult to obtain, and could increase the regulatory burden. Ofwat would need to explore, in conjunction with the regulated companies the scope for (independently verified) input prices.
Heterogeneity	Heterogeneity at this aggregated level will be less easily identifiable in relation to its impact on costs (i.e. some heterogeneity across business units will be 'averaged out' at this level). However, key operating differences like density of connections can be identified. Intertemporal differences related to operating characteristics such as levels of service might need to be controlled for	When possible, Ofwat should control for heterogeneity, through the use of appropriate variables and controls that capture statistically significant differences at this level of operation. Ofwat may need to collect additional data in order to include controls for the heterogeneity.
Opex-capex	Opex-capex trade-off will be a key consideration at this level.	Economic cost modelling states that opex modelling could include controls for changes in capital stocks. Ofwat should, where possible, include controls for changes in capital stocks (e.g. quasi-fixed capital stocks) in its opex modelling. Totex modelling is an alternative for Ofwat to use, however we consider this to be slightly less preferable as there are definitional issues in relation to what should be

[49] This was discussed further in Section 3.3.

		included (e.g. does it include capex, or depreciation and return on planned capex).
Other factors		
Technical change	At this level it is likely that technical change will impact on the performance measure.	Ofwat could include variables in the model to control/ allow for technical change. At this level the panel data length should be sufficient to allow for technical change to be controlled for. For PR14 pooled data over three to four years may be an option for Ofwat, accepting that the level of technical change is limited.
Cost interactions	Vertical cost interaction issues are likely to be relatively minor at this level of aggregation. Some potentially horizontal cost interactions may exist between *water*, *sewerage* and *retail*.	Ofwat could assess horizontal cost interactions by carrying out a consistency check at the company level. If interactions exist, then Ofwat can control for them through additional explanatory variables or through regulatory discretion.
Scale	*Water* companies may not be at the optimal size for the industry, meaning some may be experiencing economies of scale while others, diseconomies.	Ofwat could use 'flexible models' that allow for variable returns.
Length and breadth of panel	Panel data would have been collected since 2009/10 and there is some potential for the possible use of data at this level from before this period.[50] Note, data definitions can change and explanatory variables can be updated over time resulting in a reduction in data consistency.	There should be sufficient stable observations available at this level of aggregation to allow Ofwat to undertake robust performance measurement. There should also be sufficient observations to allow Ofwat to include controls for the some of the factors identified (e.g. cost interactions, heterogeneity) if they are required. Ofwat may also need to control for intertemporal changes in data (i.e. including dummy variables to take into account structural breaks).

Vertical activity level (i.e. sub-company)

The vertical segments considered under Ofwat's existing accounting separation guidelines are: *water resources*, *raw water distribution*, *water treatment* and *treated water distribution*. At this level the performance measure can provide information on cost drivers and assist in determining efficient costs, but it is not necessarily a requirement if the price control is not applied at this level.

At PR14 it is unlikely that Ofwat would need the performance measure to facilitate competition at this level, but it may provide a further check (in addition to the water level performance assessment) of cost allocation consistency if performance measurement is carried out on sub-activities. However, beyond PR14 a performance measure may be required to facilitate competition for *water resource* and possibly *water treatment* activity levels and there may be value in the early establishment of these models, and by extension data requirements, for these segments.

[50] Note, mergers will reduce the number of comparators, and the affect on the temporal element would need to be controlled for.

As with the 'total' *water* level, there is likely to be a stable panel dataset available but the length of the data set may be more appropriate for facilitating pooled panel models and greater reliance on more advanced panel models may need to be set off until future price reviews (beyond PR14) when longer panel data sets could be available.

Broadly, the key considerations for performance measurement at this level are the same as outlined for *water* in Table 6.1, i.e. controls for input price differentials, heterogeneity, technical change, etc, are required. However, some key considerations for Ofwat to consider if it were to conduct performance measurement at this level are:

- Ofwat should include capital stock measures to take account of the opex-capex trade-off.
- Cost interactions between vertical segments are likely to be high. In practice water supply costs are highly interactive through the water supply value chain. For example, companies will aim to minimise pumping and treatment costs so the cost experienced at zonal level is directly related to the availability of the source, raw water distribution, and treatment plant.
- Scale issues are still likely to exist at this level as the vertical segments should reflect the size of *water* for each company. Ofwat could use flexible models that allow for variable returns.
- There should be sufficient stable accounting data available at this level of aggregation to undertake robust performance measurement. However, Ofwat may need to collect additional information on input and output explanatory variable in order to control for factors outlined above.

Table 6.2 below provides more specific considerations for cost modelling in each of the water vertical segments. As the reporting requirements have only recently been introduced (2009/10) and the costs for these segments have yet to be analysed, the considerations presented in the table are our assumptions and are intended to illustrate the potential considerations. As with any cost modelling, Ofwat would need to analyse and test the data to determine the characteristics of each of the vertical segments.

We note that, at this level of the value chain any performance measurement methodology is feasible in principle. However, the choice of methodology is dependent on Ofwat's objectives. As discussed in Section 3.3 econometric approaches provide the greatest scope for controlling for heterogeneity, but more advanced DEA techniques also allow for some degree of heterogeneity. Both methods could be used to provide Ofwat with additional information. Non-parametric index number techniques are less plausible at this level given the inflexibility in controlling for different operating environments and the requirement for output prices proxies, which may not exist at PR14. If Ofwat requests data on variables potentially required to control for heterogeneity, scale, etc, the regulatory burden on companies would increase (unless already received as part of the June reporting requirements) and so Ofwat should consider this as part of their cost benefit analysis.

Table 6.2: Water vertical segments (i.e. business units) - summary table of segment specific key considerations

Issue	**Key considerations for each of the Water business units (at PR14)**			
	Water resources	**Raw water distribution**	**Water treatment**	**Treated water distribution**
Heterogeneity	Heterogeneity will exist between the different companies (e.g. borehole versus reservoir). Ofwat would need to undertake tests to check the significance of the heterogeneity and allow for it if necessary. [51]	Heterogeneity is likely to be driven by the source of the water, i.e. type of water (groundwater, upland river, lowland river, upland reservoir, lowland reservoir, brackish, sea water), the distance and terrain to the treatment plant.	Significant heterogeneity is likely between different treatment technologies. The cost of treat-ment is dependent on the source e.g. borehole or upland reservoir supplies are often cheaper to treat than brackish water.[52]	There is likely to be reasonable heterogeneity across companies due to different operating characteristics such as the distance from the treatment plant and the density of customer connections.
Cost interactions	Vertical cost interactions are likely, particularly with *water treatment* and *water distribution*. Ofwat should test for cost interactions and, if required, make appropriate controls or an adjustment through regulatory discretion.	Vertical cost interactions are possible. Ofwat should test for cost interactions and include appropriate controls or an adjustment through regulatory discretion. These may differ for different types of water sources.	Vertical cost interactions are likely to exist, particularly with *water resources*. Ofwat should test for cost interactions and include appropriate controls or an adjustment through regulatory discretion.	Vertical cost interactions between *water resource* and *water treatment* are likely and particularly for boreholes. Ofwat should test for cost interactions and include appropriate controls or an adjustment through regulatory discretion.
Scale	Scale issues are likely to exist for water resources. There may be scale aspects in relation to the different type of sources (e.g. large reservoirs may have different economies compared to small reservoirs). Similarly efficient scale may vary for boreholes and reservoirs.	We consider that there are likely to be scale issues for *raw water distribution* as a result of different network lengths across companies.	Scale issues are likely to exist for *water treatment*, e.g. different economies of scale between large and small treatment plants, and these are influenced by source of water being treated.	Scale issues are likely to exist for *treated water distribution*, e.g. there are likely to be variable returns to scale with at first increasing and then decreasing returns to scale as size increases, but these scale economies are also influenced by settlement patterns as larger settlements will have a larger optimal scale.

[51] For example, Ofwat allowed for heterogeneity in its PR09 model for 'Water resources and treatment' by including the number of sources divided by distribution input (to capture average source size).

[52] For example, Ofwat allowed for heterogeneity in its PR09 model for 'Water resources and treatment' by including the proportion of supplies from boreholes.

Sub-activity level

As noted above, we have based the sub-vertical activity units on possible DMUs at lower levels of the value chain. A performance measure for PR14 is plausible at this level of the value chain and it may be required for determining efficient costs – especially to support greater competition. However, as with the vertical segments, we recommend that Ofwat should undertake a cost-benefit analysis (Section 8) considering the additional data requirements, accuracy issues and resource requirements against the additional information provided by the performance measurement.

If reporting requirements are put in place soon, it is plausible that two or more years of data may be available for PR14, but with caveats around the stability and accuracy. For PR14, one year of data may be sufficient for performance measurement, however, this is dependent on the number of comparators available and the number of variables required for the models. Therefore, Ofwat could begin collecting this information, but it should consider the modelling requirements before it can commit to undertaking performance measurement at this level for PR14.

The key considerations for performance measurement of sub-activity units in *water* are similar to those set out for the *water* vertical segments. Some key issues that Ofwat would need to consider to be able to carry out performance measurement at this level include:

- Heterogeneity will be more apparent at this level of analysis and Ofwat will almost certainly need to use some controls for differences identified. For example, the *water resource* sub-activity is likely to have different operating characteristics within and across the companies.
- The ability to trade-off opex for capex will exist at this level e.g. introduction of remote monitoring at a treatment plant as a substitute for labour. Ofwat should introduce controls for capital, however at this level it may be more difficult (or a least require further information) to establish the asset base for each sub-activity unit.
- There will likely be significant difference in scale which may result in biased estimates if (dis)economies of scale are not taken into account. Where possible Ofwat should use controls for different sized DMUs, this may be through additional explanatory variables or different models for groups of different sizes. And
- Data is not currently being collected at this level. Given the time required to develop and consult on reporting guidelines, it is unlikely that more than two years of panel data would be available by PR14. There is likely to be a reasonably high additional cost, to both Ofwat (processing the data and ensuring consistency) and the regulated companies, to collecting the required variables at this level.

We do not present a more detailed breakdown of specific considerations for each of the sub-activities as further analysis would need to be undertaken to inform any consideration of cost modelling at this level.

6.1.3. Summary

Given the accounting separation guidelines set by Ofwat, at PR14 a stable but relatively short panel dataset (three to four years) would be available for conducting performance assessment at the potential regulated business (*water*) and vertical segment levels (*water resources*, *raw water distribution*, *water treatment* and *treated water distribution*). Given the potential cost interactions between sewerage and water activities for WaSCs, company level consistency checks are important to determine if economies of scope exist and whether these need to be controlled for. Analysis done at the potential regulated business level will be required for the price control.

As cost data is currently collected at the vertical segment (sub-company) level, performance measurement by Ofwat at this level will likely result in relatively low levels of additional resource cost, e.g. collection of output and input explanatory variables. Issues around accurate cost allocation, vertical integration economies, and (dis)economies of scale would need to be considered as part of a cost benefit analysis, however it is likely that performance measures at this level will provide significant information on cost drivers.

Undertaking performance measurement at the sub-activity level would appear to offer few benefits as it would not be required for facilitating competition or consistency checks at PR14. In addition, given the shorter 'settling' period for the collection of the data from the regulated companies there is a risk that the data would be less robust (in relation to cost allocation) and stable.

The conclusions of this section as they affect PR14 are set out in Tables 6.3 and 6.4 below.

Table 6.3: Evaluation of performance measurement feasibility at PR14

Activity level	**Sub-activity level**	**Performance measurement requirement**			**Continue?**
		Efficient cost determination	**Facilitating efficient competition**	**Consistency check**	
Water	-	Yes	No	Yes	Yes
Water resources	-	Yes	No	*Yes	Yes
Water resources	Zone	**Yes	**Yes	No	**
Water resources	Source	**Yes	No	No	**
Water resources	Trade	**Yes	**Yes	No	**
Raw water distribution	-	Yes	**Yes	No	Yes
Water treatment	-	Yes	**Yes	*Yes	Yes
Water treatment	Trade	**Yes	**Yes	No	**
Treated water distribution	-	Yes	**Yes	*Yes	Yes
Treated water distribution	WS Zone	**Yes	**Yes	No	**
Treated water distribution	Sub business operating unit	**Yes	**Yes	No	**

* Consistency check required if benchmarking performed at sub-vertical activity level.

** Theoretically feasible, but dependent on Ofwat's requirements/ objectives Facilitating competition is unlikely at PR14, however potentially needed beyond PR14 and is discussed further in Section 7.

Table 6.4: Evaluation of performance measurement feasibility at PR14

Activity level	Sub-vertical activity level	Data availability (years)[1]	Methodology			Continue?
			Econometrics	DEA	Non-parametric index[2]	
Water	-	<4	Yes	Yes	Yes	Yes
Water resources	-	<4	Yes	Yes	No	Yes
Water resources	Zone	<2	Yes	Yes	No	*
Water resources	Source	<2	Yes	Yes	No	*
Water resources	Trade	<2	Yes	Yes	No	*
Raw water distribution	-	<4	Yes	Yes	No	Yes
Water treatment	-	<4	Yes	Yes	No	Yes
Water treatment	Trade	<2	Yes	Yes	No	*
Treated water distribution	-	<4	Yes	Yes	No	Yes
Treated water distribution	WS Zone	<2	Yes	Yes	No	*
Treated water distribution	Sub business operating unit	<2	Yes	Yes	No	*

[1] Based on the assumption of only using historical data for benchmarking.

[2] Non-parametric indices are feasible at lower levels of the value chain, however we consider that they are less desirable at lower levels (see Section 3.3.3 for a further discussion around non-parametric indices).

* Feasible, but dependent on the number of observations available.

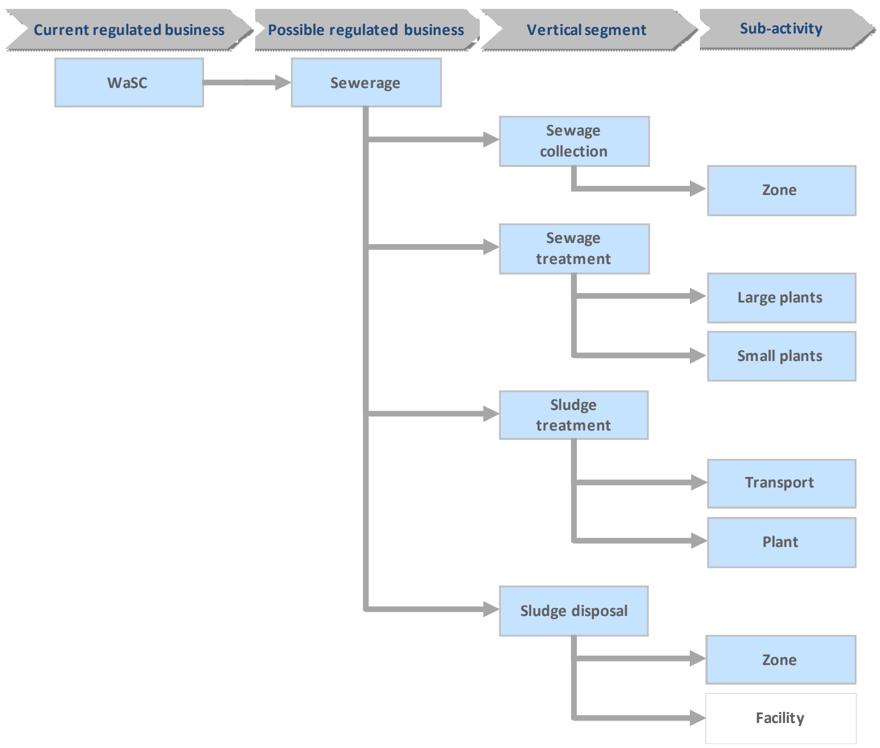
6.2. Sewerage

As with *water* above, we consider the use of accounting separation data is a logical extension of Ofwat's current modelling approach.

6.2.1. Possible structure

Figure 6.2 provides an illustration of the possible structure of the sewerage sector under the current accounting separation guidelines. As with the *water* sector the vertical segments are based on Ofwat's business units, while the sub-activity level illustrates plausible (not necessarily probable) smaller DMUs. Note, the *sewerage* structure excludes *retail*, which is discussed in Section 6.3.

Figure 6.2: Possible structure of 'Sewerage'



6.2.2. Evaluation

As with Water supply, we have grouped the different DMUs at similar levels across the business in order to restrict the number of independent assessments.

Sewerage

We assume that the price control is applied at this level, therefore a performance measure is required for efficient cost determination and potentially required consistency checks (for cost allocation issues if analysis is undertaken below this level of aggregation). The performance measure is not required to facilitate competition as this is infeasible at this level of the value chain for PR14.

As with *water* there is likely to be a stable three to four year panel dataset available as: (i) the accounting separation guidelines have been in place since the 2009/10 reporting year; and (ii) as this is a high level of the value chain (e.g. high level of cost aggregation), cost allocation issues will be relatively minor.

The key factors that Ofwat needs to consider in relation to Sewerage are very similar to those in Water (please refer to Table 6.1). Briefly, they are:

- The potential need to include variables in Ofwat's cost modelling to control for various factors (e.g. heterogeneity) at this level, this may require Ofwat to use panel or pooled data in order to provide sufficient degrees of freedom.
- Sewerage companies may not be at the optimal size for the industry, meaning some may be experiencing economies of scale while others, diseconomies. Ofwat could use flexible models that allow for variable returns to scale for different companies.
- Geographically driven input price differentials are likely to exist. Ofwat would need to consider whether there is a requirement to make these adjustments and, if so, the appropriate data and adjustment to use, e.g. ONS.
- Heterogeneity will exist across companies, but at this level it is unlikely variables could be included to sufficiently capture this (i.e. as the aggregated costs include costs from all Sewerage vertical segments, controlling for differences in each at this level would be difficult). Given the change in outputs that has been occurring in sewerage, this is likely to be an important consideration – especially for treatment. And
- While vertical cost interactions issues are likely to be relatively minor at this level of aggregation, some potential horizontal cost interactions will exist, particularly with water in the case of WaSCs. Ofwat could assess horizontal cost interactions through a consistency check at the company level. If interactions exist Ofwat could control for these or take them into account through regulatory discretion.

At this level of the value chain any performance measurement methodology is feasible. There are, however, relatively fewer comparators in Sewerage (10, in comparison to the 21 in Water) meaning that there would be a smaller panel data set available for cost modelling than in Water, i.e. Sewerage would have fewer observations and hence degrees of freedom than for Water (40 observations available in a four year panel data set). This raises concerns that there may be insufficient observations at the company or sub-company level for robust estimates to be produced.

We believe that there should be a sufficient number of observations to allow for additional variables to take account of, for example, technical change (if tests indicated that these variables are required) – although this may restrict the modelling techniques available (e.g. SFA may not be possible). Alternatively, Ofwat can use pooled data rather than panel, if tests showed that technical change did not need to be controlled for. If the panel was shorter (e.g. less than four years) Ofwat would need to consider whether the number of observations were sufficient to provide robust estimates. We note that there is no definitive rule in regards to the minimum number of degrees of freedom required, however we note that: (i) Ofwat undertook regressions using only 22 observations for PR09; and (ii) Ofgem carried out modelling with panel data on group ownership for DPCR5 rather than company ownership, which meant only 28 observations (seven comparators over four years) were available. These examples indicate that regulators consider that relatively fewer numbers of observations can produce useable estimates. Of course a greater number of observations are always preferred in order to increase confidence in the results.

While we believe that there will be sufficient observations available, we have not conducted tests regarding the degree of heterogeneity and technical change in Sewerage. Ofwat would need to conduct these tests before determining the number of variables required, and hence whether there are sufficient degrees of freedom to: (a) run a correctly specified model; and (b) produce robust estimates.

Vertical activity level (sub-company)

The vertical segments considered under the accounting separation guidelines are: *sewage collection*; *sewage treatment*; *sludge treatment*; and *sludge disposal*. At this level the performance measure can provide information on cost drivers and assist in determining efficient costs. Ofwat has previously investigated modelling sludge using econometric (or DEA) approaches, however given the multiple alternative uses of sludge and the degree to which this is beyond the managerial control this has so far been unsuccessful. For example, the cost structure for incineration of sludge is quite different from agricultural use. Ofwat has, however, used a unit cost model for 'sludge treatment and disposal'. The water industry is currently producing a report looking at modelling sludge opex efficiency and Ofwat is on the steering group.

It is important to note that policymakers may determine that competition should be introduced for *sludge treatment* and *sludge disposal*, as a recent report commissioned by Ofwat noted.[53] Ofwat has recently announced a joint study with OFT into the treatment of organic waste.[54] Therefore, given the potential scope for competition in Sludge treatment and disposal this is an area where performance measures for facilitating competition may be required, possibly at PR14 but more likely beyond PR14.

As with the 'total' *sewerage* level, there is likely to be a stable panel dataset available given the existing accounting separation guidelines. However, greater reliance on more advanced panel models would need to be held off until future price reviews when longer panel datasets would be available.

The considerations for the use of performance measures at the *sewerage* vertical segment level are very similar to those for *water* vertical segments and *sewerage* (outlined in the preceding section). Table 6.5 below provides some specific considerations for cost modelling for each of the Sewerage vertical segments.

---

[53] London Economics (2010).

[54] Ofwat press release, 21 February 2011.

Table 6.5: Sewerage vertical segments - summary table of segment specific key considerations

Issue	**Key considerations for each of the Sewerage business units (at PR14)**			
	Sewage collection	**Sewage treatment**	**Sludge treatment**	**Sludge disposal**
Heterogeneity	There is likely to be reasonable heterogeneity across companies due to different operating characteristics such as the topography, rural/urban split, and density of customer connections.[55] Other local factors may also be important. For example infiltration of ground water to sewers may have a significant effect on pumping costs in some catchments. In catchments with relatively large non-domestic flows, input costs may be dominated by trade effluent.	There is likely to be reasonable heterogeneity across companies i.e. effluent consents vary by area (more expensive to meet tighter effluent consents).[56] In addition to the variation introduced by discharge consents, the presence of specific trade effluents may affect the choice of treatment process and hence the capital and operating costs of the treatment plant.	There is likely to be reasonable heterogeneity across companies, driven by the distance to transport sludge to the treatment plant and the different approaches to treating it.	There is significant heterogeneity between different companies in relation to the routes to disposing of treated sludge e.g. incineration has quite different costs from agricultural use.
Cost interactions	Vertical cost interactions are possible for this segment. Ofwat should test for cost interactions and, where required, adjustments made.	Vertical cost interactions are possible for this segment. Ofwat should test for cost interactions and, where required, adjustments made. In particular there will be interactions between trade and domestic effluent treatment costs that may be difficult to separate, because the balance and extent	There are cost interactions between *sewage treatment*, *sludge treatment* and *sludge disposal* (for instance, dewatering). Large sewage treatment plants are likely to have adjacent sludge treatment facilities, providing economies of land, transport and management. Ofwat should control for these	Vertical cost interactions are possible for this segment. The distance to 'market' for the disposal route, and choice of technology, will have a significant effect on cost. In some cases the choice of disposal route is limited, e.g. some urban sludge is not suitable for agricultural use. This may provide

55 For example, Ofwat allowed for heterogeneity in its PR09 model for 'Network including power' by including variables taking into account density of the network.

56 For example, Ofwat allowed for heterogeneity in its PR09 model for 'Large sewage treatment works' by including variables taking into account the nature of the effluent consents.

		may differ on a site-by-site basis. Ofwat may need to make an assessment of trade effluent effects and adjusted for these if required.	factors , or make allowance for them in the regulatory discretion decision.	a further interaction with the trade effluent effects on collection and treatment.
Scale	Economies of scale are likely to exist, and Ofwat should therefore test for and, where required, adjusted for this.[57]	Economies of scale are likely to be significant, and Ofwat should therefore test for and, where required, adjusted for this.	Economies of scale are likely to exist at the plant level if there is sufficient feedstock within economic transport distance.[58] Ofwat would need to test for these and, where required, adjust for.	Economies of scale are likely to exist, but subject to transport distance. Ofwat would need to test for this and, where required, adjusted for.

[57] For example, Ofwat allowed for scale in its PR09 model for 'Network including power' by including variables taking into account population serviced and network length.

[58] Feedstock refers to the material used as the main input to the sludge treatment process. For example, in the case of anaerobic digestion, this will consist of organic material such as sewage sludge, green waste, food processing waste and some industrial process waste. When selecting feedstock operators will also consider: whether it contains materials that may affect handling within the process; its impact on digestion of other feedstock; and/or the suitability of the digestate for the preferred final disposal route.

As with *water* we note that, at this level of the value chain any performance measurement methodology is feasible. However, the choice of methodology is dependent on Ofwat's objectives. As discussed in Section 3.3 econometric approaches provide the greatest scope for controlling for heterogeneity, but more advanced DEA techniques also allow for some degree of heterogeneity. Both techniques could be used to provide Ofwat with additional information. Non-parametric index number techniques are less desirable at this level given that they are less flexible when controlling for different operating environments and the requirement of output price proxies (which may not be available at PR14). If Ofwat requests data on variables potentially required to control for heterogeneity, scale, etc, the regulatory burden on companies would increase (unless already received as part of the June reporting requirements) and so Ofwat should consider this as part of their cost benefit analysis.

Sub-activity level

As noted above, we have based the sub-activity units on possible DMUs at lower levels of the value chain. At this level of the value chain a performance measure for sewerage at PR14 is only plausible for determining efficient costs. However, given the potential scope for competition in sludge performance measures for *sludge treatment* 'transport' and 'plant' and *sludge disposal* 'zone' and 'facility' may be required to determine efficient access prices. Separate models for different plant types (etc) may be more suitable given the level of heterogeneity; however this is dependent on available observations and controlling for environmental differences. The objectives of performance measures need to be clear, whether it is:

- to increase information available to new entrants;
- to determine more accurate access pricing; and/or
- as an ongoing performance measure.

If Ofwat puts accounting requirements in place soon it is plausible that two or more years of data may be available, but with caveats around the stability and accuracy. Given the potential importance of this information for competition, significant work will need to go into ensuring that the cost allocation is correct.

DEA methodologies are feasible at this level of the value chain. However, given their greater ability to cope with heterogeneity between comparators, econometric techniques would probably be superior. But, the difficulties experienced in developing cost models for sludge suggest that, even though models at this level are theoretically feasible, they may not be practical.

The key considerations for performance measurement of sub-activity units for *sewerage* are similar to those detailed for *water* sub-activities, however we note some examples of issue for Ofwat to consider below:

- Heterogeneity will likely be apparent at this level of analysis. For example, treatment plants needing to comply with different effluent consents. In addition, the trade effluent that plants have to deal with is also very variable (e.g. dairies, ice-cream makers and meat processors produce very strong effluent whereas others may not).

- There will likely be significant difference in the *scale* which may result in biased estimates if (dis)economies of scale are not taken into account. Large sewage treatment plants generally have lower per output operating costs compared to smaller plants.

6.2.3. Summary

While on the whole our view of the use of panel and sub-company data is similar for water and sewerage, it is clear that one area where facilitating competition places a greater reliance on sub-company data within *sewerage* is *sludge treatment* and *sludge disposal*. This raises a number of issues particularly given Ofwat's inability so far to estimate stable and robust econometric models for these activities.

The conclusions for the Section as they affect PR14 are set Table 6.6 and 6.7.

Table 6.6: Evaluation of performance measurement feasibility at PR14

Activity level	Sub-vertical activity level	Performance measurement requirement			Continue?
		Efficient cost determination	Facilitating efficient competition	Consistency check	
Total Sewerage	-	Yes	No	Yes	Yes
Sewage collection	-	Yes	No	*Yes	Yes
Sewage collection	Zone	**Yes	No	No	**
Sewage treatment	-	Yes	No	*Yes	**
Sewage treatment	Large plants	**Yes	No	No	**
Sewage treatment	Small plants	**Yes	No	No	Yes
Sludge treatment	-	**No	**Yes	*Yes	Yes
Sludge treatment	Transport	**No	**Yes	No	**
Sludge treatment	Plant	**No	**Yes	No	Yes
Sludge disposal	-	**No	**Yes	*Yes	Yes
Sludge disposal	Zone	**No	**Yes	No	**
Sludge disposal	Facility	**No	**Yes	No	**

* Consistency check required if benchmarking performed at sub-vertical activity level

** Theoretically feasible, however the level of heterogeneity and but dependent on Ofwat's requirements/ objectives. Facilitating competition is unlikely at PR14, however potentially needed beyond PR14 and is discussed further in Section 7.

Table 6.7: Evaluation of performance measurement feasibility at PR14

Activity level	**Sub-vertical activity level**	**Data availability[1] (years)**	**Methodology**			**Continue?**
			Econometrics	**DEA**	**Non-parametric index[2]**	
Total Sewerage	-	<4	Yes	Yes	Yes	Yes
Sewage collection	-	<4	Yes	Yes	No	Yes
Sewage collection	Zone	<2	Yes	Yes	No	*
Sewage treatment	-	<2	Yes	Yes	No	*
Sewage treatment	Large plants	<2	Yes	Yes	No	*
Sewage treatment	Small plants	<4	Yes	Yes	No	Yes
Sludge treatment	-	<4	**Yes	**Yes	No	Yes
Sludge treatment	Transport	<2	**Yes	**Yes	No	*
Sludge treatment	Plant	<4	**Yes	**Yes	No	Yes
Sludge disposal	-	<2	**Yes	**Yes	No	*
Sludge disposal	Zone	<2	**Yes	**Yes	No	*
Sludge disposal	Facility	<2	**Yes	**Yes	No	*

[1] Based on the assumption of only using historical data for benchmarking.

[2] Non-parametric indices are feasible at lower levels of the value chain, however we consider that they are less desirable at lower levels (see Section 3.3.3 for a further discussion around non-parametric indices)

* Feasible, but dependent on the number of observations available.

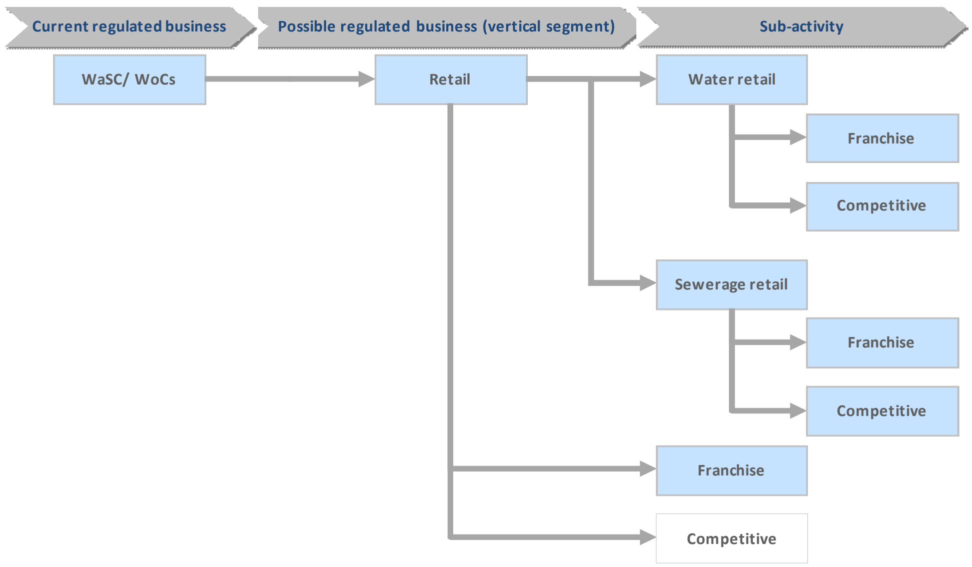
** Theoretically feasible, however Ofwat has attempted to model this and the degree of heterogeneity has prevented robust models from being developed. Also the choice of plant is to some extent outside the control of management.

6.3. Retail

6.3.1. Possible structure (Retail)

In this section, as per Figure 6.3, we are focusing on the third and final aspect of the accounting separation guidelines split, that of retail. Ofwat's existing accounting separation guidelines only require one *retail* business unit's to be reported, i.e. Ofwat does not require water and sewerage retail costs to be reported separately.

Figure 6.3: Possible structure of 'Retail'



When thinking about retail it is possible to see a few ways in which sub-activities could be identified. These include a split between:

- Water and sewerage – given the existence of WoCs, some WaSCs have different coverage for their water retailing and sewage retailing (although this may often be addressed through existing contractual relationships between the WoC and WaSC). And
- Competitive and franchise (non-competitive) customers – the latter being primarily households where it is not envisaged (at least in the Cave Review) to introduce competition in the short- to medium-term.

6.3.2. Evaluation

We consider below which uses of measurement could usefully be applied to *retail* at the regulated and vertical levels.

Retail

This is the assumed level that the price control is applied i.e. to a retail segment comprising water (WoCs) or water and sewerage (WaSCs). At this level a performance measure is required for efficient cost determination and consistency check (for cost allocation issues). The performance measure is not required to facilitate competition as we consider that it is infeasible for this to occur at this aggregated level of the value chain by PR14 as household competition is very unlikely and how competition for non-household customers will develop is very dependent on the forthcoming Water White Paper.

For PR14, there is likely to be a stable three to four year panel dataset available as: (i) the accounting separation guidelines have been in place since the 2009/10 reporting year; and (ii) as this is a high level of the value chain (e.g. high level of cost aggregation), cost allocation issues will be relatively minor.

At this level of the value chain any performance measurement methodology is feasible. However, the choice of methodology is dependent on Ofwat's objectives.

We consider that separate analysis of water retail and sewerage retail should be undertaken by Ofwat since the WoCs will only be providing the former while WaSCs provide both. Or at least these potential cost interactions and economies of scale should be controlled for. Note, this split is not set out in the accounting separation guidelines, but it could be controlled for using a explanatory factor. The implications of this are discussed under sub-company analysis below.

Table 6.8: Key considerations for performance measurement for retail

Issue	Relevance to this part of the value chain	Implication and solutions
Underlying economic theory		
Input costs	Geographical differences likely to be the main driver of input differentials. However, these may be less than for the water and sewage businesses as all retail activities do not need to be located in the area of water or sewage supply.	At this level it is likely that input price differentials will lead to biased estimates. Ofwat could control for these with the introduction of additional variables. However, it is not clear that these should be controlled for. As, especially over the medium-term, inputs to Retail are less prone to significant price differentiation (since they can be sourced nationally, such as a call centre) and industry reform may also change the nature of the businesses. A proportion of cost may be mobile and companies could control some of the costs through a move to lower input price areas.
Heterogeneity	There is likely to be relatively low heterogeneity given the nature of the outputs.	If required, Ofwat could introduce variables to control for heterogeneity between the different companies.
Opex-capex	The ability to trade-off opex for capex will exist at this level. However, it is likely these business units will be 'asset light' and consequently the scope for opex capex trade-off will be limited.	Given the likely low proportional levels of capex to opex, allowing for a trade-off may not be needed. Ofwat should test the scope for trade-off within these business units.
Other factors		
Technical change	Technical change will potentially exist for retail units, for example, newer billing techniques, etc.	Where possible Ofwat should control for technical change if tests show it to be present in the data set.
Cost	Vertical cost interactions are less	Cost interactions can be assessed by Ofwat at a

interactions	likely for this business unit (than in Water or Sewerage), however horizontal cost interactions will need to be considered.	higher level and Ofwat could introduce controls into the modelling if interactions are found to exist, or alternatively these could be dealt with through regulatory discretion.
Scale	Scale issues may exist for this business unit.	It is not clear that this is a concern for *retail*, especially as outsourcing/joint working already happen. Ofwat would require a better understanding of how the retail function is going to work in the future to provide a definitive answer to whether scale is (or should be) a concern. It is likely that economies of scale would exist for this type of services and Ofwat will need to decide whether to include this in its models.
Length and breadth of panel	Panel data would have been collected since 2009/10.	We consider that there should be sufficient stable observations available at this level of aggregation to undertake robust performance measurement.

Sub-activity level

As noted above, we consider two separate sub-activity splits for *retail:* (i) a vertical split into *water retail* and *sewerage retail*; and (ii) a horizontal split into f*ranchise* and c*ompetitive.*

For the water/ sewerage split, the key considerations are similar to those set out in Table 6.8 above, Given the likely horizontal cost interactions between the segments (for WaSCs), Ofwat should employ economies of scale consistency checks to test for biased results. If checks indicate economies of scale and/ or scope then the performance measure should control for these.

For *franchise* and *competitive*, the performance measure can provide information on cost drivers and assist in determining efficient costs and also facilitate competition for the customers within the Competitive segment. If contestability in the Competitive segment was sufficient then performance measurement may not be required. Given the likely cost interactions, economies of scale and scope unless consistency checks are employed biased results are likely.

6.3.3. Summary

While in some respects *retail* may be the simplest of the activities, the uncertainty about whether the two vertical activities will be considered as separate activities for sub-price caps or performance measurement does create some issues. It is clear that panel data offers a feasible way of assessing performance at PR14. It is also likely that performance measurement at a lower level will be needed either for price setting for the WoCs or to facilitate competition for the competitive customer segment of the market. As such, it will be important to undertake consistency checks so that any bias in the results can be addressed.

Table 6.5: Evaluation of performance measurement feasibility at PR14

Activity level	**Sub-vertical activity level**	**Performance measurement requirement**			**Continue?**
		Efficient cost determination	**Facilitating efficient competition**	**Consistency check**	
Retail	-	Yes	Yes	Yes	Yes
Water Retail	-	**Yes	Yes	*Yes	Yes
Water Retail	Franchise	**Yes	No	No	Yes
Water Retail	Competitive	No	Yes	*Yes	**
Sewerage Retail	-	**Yes	Yes	*Yes	Yes
Sewerage Retail	Franchise	**Yes	No	No	Yes
Sewerage Retail	Competitive	No	Yes	*Yes	**
Franchise	-	**Yes	No	No	Yes
Competitive	-	No	Yes	*Yes	**

* Consistency check required if benchmarking performed at sub-vertical activity level

** Feasible, but dependent on Ofwat's requirements/ objectives.

Table 6.6: Evaluation of performance measurement feasibility at PR14

Activity level	Sub-vertical activity level	Data availability[1] (years)	Methodology			Continue?
			Econometrics	DEA	Non-parametric index	
Retail	-	<4	Yes	Yes	Yes	Yes
Water Retail	-	<4	Yes	Yes	No	Yes*
Water Retail	Franchise	<2	Yes	Yes	No	Possibly
Water Retail	Competitive	<2	Yes	Yes	No	Possibly
Sewerage Retail	-	<4	Yes	Yes	No	Yes*
Sewerage Retail	Franchise	<2	Yes	Yes	No	Possibly
Sewerage Retail	Competitive	<2	Yes	Yes	No	Possibly
Franchise	-	<2	Yes	Yes	No	Possibly
Competitive	-	<2	Yes	Yes	No	Possibly

[1] Based on the assumption of only using historical data for benchmarking.

[2] Non-parametric indices are feasible at lower levels of the value chain, however we consider that they are less desirable at lower levels (see Section 3.3.3 for a further discussion around non-parametric indices)

* Feasible, but dependent on the number of observations available.

6.4. Consistency check

An important aspect of undertaking analysis based on data reported in line with the accounting separation guidelines is the need to undertake a consistency check at the (vertical segment) ownership level. It is widely acknowledged that undertaking bottom-up assessment does not necessarily capture the full scope for efficiency gains that a company can achieve (for example, if cost interactions exist between segments/ activities then biased results would be found). This will become a key consistency check for the overall performance targets being set at PR14 and beyond.

7. BEYOND PR14

Key themes and lessons

What happens beyond PR14 is much less certain since the degree of competition in elements of the industry may well be greater and, consequently, there may be less need for performance measurement for price-setting but possibly more needed for facilitating and supporting competition. Further industry structure changes may also take place, such as separation of system operators.

It is likely that Ofwat will need to consider both panel and sub-company data beyond the next price review. The precise need will depend on the final industry structure and the way that competition develops. However, what is clear is that some elements of the industry, especially related to the pipes and other essential facilities, will require ongoing conduct regulation. Where conduct regulation is needed price determinations will be needed which in turn requires cost assessments and performance measurement.

As competition develops in some elements of the value chain, both for water and sewerage, requirements for greater cost transparency and localised access prices may be needed. If this is the case a role for performance measurement may arise, especially at the sub-company and sub-activity levels.

There will be much greater scope for the use of more sophisticated panel data methods beyond PR14 as the length of relevant data series could be up to five years longer. This in itself would help improve estimates derived from panel data models, including their robustness.

7.1. Introduction

There is much greater uncertainty about both industry structure and the focus of regulatory price controls beyond PR14. In part the possibilities for the future will be guided by the Water White Paper due from Defra later this year. New industry structures may arise as greater competition in areas like water resources and sludge treatment/disposal create new opportunities and requirements for the sector. Consequently we need to consider performance measurement in two areas:

- cost assessment for aspects of the sector that require conduct regulation; and
- facilitating new entrants through greater cost transparency and, where necessary, local access charge in the contestable aspects of the sector.

This section considers ways in which the industry may develop beyond PR14 and builds on elements of the recent Ofwat publication *Future price limits – a preliminary model (informal consultation)*.[59] However, the lack of precision with respect to future industry structures means that the discussions in this section are high level.

7.2. Possible future structures

As shown in Section 6, Ofwat's accounting separation guidelines provide a greater scope for applying price caps at lower 'functional' levels of a company's structure. Section 6 set out three plausible price caps levels, *water, sewerage* and *retail*, to illustrate our approach for determining the

[59] Ofwat (2011a).

use of performance measure at these levels. However, there are many possibilities for regulating the sector in the future. Ofwat's recent publication sets out a preliminary model of how it might set price controls in the future. The preliminary model is based on Ofwat's assumption that the structure of regulation should reflect the different characteristics of the activities in the water and sewerage value chain. Ofwat's proposed preliminarily model outlines three main business units, two wholesale ('network plus' and 'resources') and one retail, for which price controls could be set.[60]

For this report we need to consider how the value chains for water and sewerage might develop beyond PR14. Tables 7.1 and 7.2 consider the four elements of each of the two value chains and how structures may change and their implications for regulation. In part the comments in the tables build on the evidence from other regulated sectors. Some discussion of the implications for panel and sub-company/activity data are included in the tables below. Further consideration is provided later in this section, along with some initial thoughts on the implications for the approach to performance measurement that could be used.

Table 7.1: Possible developments in the water value chain

Element of the value chain	Possible developments	Implications for regulation and performance measurement
Water resources	Forms of competition are possible, either "for the market" or "in the market". This has been a focus of recommendations from the Cave Review. Ofwat has also considered the possible development of a system operator handling aspects of the value chain.	Regulation should become less necessary as the degree of competition develops. This should mean that the need for price determinations will also disappear and consequently there will be less need for Ofwat to undertake performance measurement in this area. There may be transition issues around information requirements depending on the type of competition. For example, if competition develops in stages then more disaggregated data may need to be collected. This would mean that Ofwat would still need to undertake performance measurement, although it may be at a sub-company/activity level. How a system operator would be regulated, or whether it needs to be regulated, also needs to be considered. For example, unless a form of competition is chosen for the system operator, such as franchising, it is likely that Ofwat would need to undertake some form of performance measurement to determine a revenue requirement.
Raw water distribution	How this element develops will depend on the way in which the raw water distribution system is treated. If it is included in the water resources then it could be subject to the development of competition. If it is treated separately then it is likely to be viewed as an essential	If included in the water resources element of the value chain, then the discussion above is relevant. If treated as an essential facility then ongoing conduct regulation will be necessary. This will imply the need for cost assessment through performance measurement.

[60] Ibid, p.19.

	facility that needs to be subject to ongoing conduct regulation.	
Water treatment	Again, this is an area where competition might develop, although initially this may be through "for the market" rather than "in the market".	If competition develops then the focus for performance measurement, at least in a transitional phase, would be for information provision. This would be to help facilitate competition and inform new entrants. As such, a focus on total costs for the sub-activity (treatment works) and their levels of efficiency could be appropriate. If competition does not develop then a need for ongoing conduct regulation will exist. Then performance measurement will need to focus on cost assessment. Ofwat's choice on the degree of aggregation for the assessment will depend on Ofwat's views of the separability of the water treatment plant costs. It will also depend on the way in which this segment is price regulated – would there be a cap for water treatment or would each plant have a separate allowed cap? Clearly sub-company/activity are both possible here, the question is whether they would be needed
Treated water distribution	This activity is a local monopoly and competition is unlikely (except for competition to build new, or replacement, distribution assets). There may be a role for performance measurement with respect to facilitating competition which requires more granular pricing depending on the way that upstream competition develops.	The local monopoly status means ongoing conduct regulation will be necessary. This will imply the need for cost assessment through performance measurement so that prices can be set. Given the likely ongoing nature of the conduct regulation panel data approaches would seem to be appropriate. Depending on the way competition develops there may be a need for local access prices to be set. These would be at a more granular level than the existing price controls. Ofwat would need to focus on sub-activity performance measurement to provide efficient cost reflective access prices.

Table 7.2: Possible developments in the sewerage value chain

Element of the value chain	Possible developments	Implications for regulation and performance measurement
Sewage collection	This activity is a local monopoly and competition in any form is unlikely. (except for competition to build new, or replacement, distribution assets).	The local monopoly status means ongoing conduct regulation will be necessary. This will imply the need for cost assessment through performance measurement so that prices can be set. Given the likely ongoing nature of the conduct regulation panel data approaches would seem to be appropriate. It is unlikely that a role with respect to facilitating competition through sub-activity analysis will exist since sewerage networks tend to be local. This may still, however, require local price determination but that would depend on the form of competition that develops.
Sewage treatment	This is an area that might be subject to the development	If competition develops then the focus for performance measurement, at least in a transitional

	of competition, although initially this may be through "for the market" rather than "in the market".	phase, would be for information provision. This would be to help facilitate competition and inform new entrants. As such, a focus on total costs for the sub-activity (treatment works) and their levels of efficiency could be appropriate. If competition does not develop then a need for ongoing conduct regulation will exist. Then performance measurement will need to focus on cost assessment. Ofwat's choice on the degree of aggregation for the assessment will depend on Ofwat's views of the separability of the sewage treatment plant costs. It will also depend on the way in which this segment is price regulated – would there be a cap for sewage treatment or would each plant have a separate allowed cap? Clearly sub-company/activity are both possible here, the question is whether they would be needed.
Sludge treatment	This is an area where competition may develop. The form of competition is unclear but the OFT is currently examining it in a market study.	If competition develops then the focus for performance measurement, at least in a transitional phase, would be for information provision. Similar issues to sewage treatment would then need to be considered. If competition does not develop then a need for ongoing conduct regulation will exist. Then performance measurement will need to focus on cost assessment. Whether competition develops or does not, similar issues to sewage treatment would need to be considered.
Sludge disposal	This is an area where competition may develop. The form of competition is unclear. If competition does not develop then it is possible that this could be an essential facility allowing upstream competition. As such, access issues might be important.	If competition develops then the focus for performance measurement, at least in a transitional phase, would be for information provision. If competition does not develop then a need for ongoing conduct regulation will exist. Then performance measurement will need to focus on cost assessment. Since sludge disposal may be important for sludge treatment it is vital to ensure that it can play a facilitating role if needed. Whether competition develops or not, similar issues to sewage treatment would need to be considered.

Retail, the third business area, is unlikely to face issues that are different to those outlined for PR14. The two options of sub-company focus – competitive/franchise and water/sewerage are likely to remain key for the period beyond PR14. Consequently the issues addressed in Section 6 with respect to PR14 are still relevant for the period beyond PR14.

This high-level analysis of what may happen beyond PR14 suggests that there are several elements of the value chain where competition could develop – although the form of competition is unclear (and there may not be a uniform approach to competition across England & Wales depending on the circumstances). However, it is also clear that some elements will remain either monopolies or essential facilities that will be subject to ongoing conduct regulation. Finally, there are some potentially new areas, like the system operator, which may require conduct regulation, although that again is uncertain.

7.3. Implications for panel, sub-company and approaches

Having considered the way in which the water and sewerage industry may develop, we believe it is clear that even with a changing industry structure there will be a need for Ofwat to undertake performance measurement. Specifically:

- conduct regulation of monopoly areas (and contestable areas where competition fails to develop) will need to continue for the foreseeable future and we expect that Ofwat will need to undertake performance measurement for cost assessment;
- where competition is developing there may be a transitional information provision role for regulation which will need to incorporate performance measurement – the length of the transition will depend on factors like the development of new facilities, degree of competition while the types of performance measurement would be focused on providing information on the costs of existing facilities and their level of efficiency; and
- depending on the form of competition that develops it is possible that local access prices will be needed for essential facilities, such as the monopoly networks, as a way of facilitating competition. If this is the case, Ofwat will need to undertake performance measurement at that local level to ensure cost reflective efficient prices are set.

Given these likely developments it is now possible to consider the implications for both the use of panel and sub-company/activity approaches as well as a more general consideration of the approach to performance measurement.

What is clear with respect to panel analysis is that:

- some areas, which require ongoing conduct regulation, are likely to be areas where it is feasible for panel data analysis to be employed since price determinations are needed and all the issues linked to the discussion in Section 6 will be relevant concerning the number of comparators available for performance assessment in these cases;
- for areas where facilitating competition could require more localised access prices then it is feasible to consider that Ofwat could use panel data analysis as a way of undertaking performance assessment since again there is likely to be a long term need for such regulation. A separate question as to whether Ofwat needs to move to more disaggregated price setting would also need to be considered at some time if this role proves to be important; and
- where transitional information provision is required to facilitate competition it is feasible for Ofwat to use panel data analysis although it would need to consider whether any costs associated with the panel data collection outweighed the short-term benefits (owing to the transitional nature).

The type of cost benefit analysis that would be necessary when considering the use of panel data is discussed in the following section.

For sub-company/activity analysis the following implications can be drawn:

- it is feasible for Ofwat to use sub-company/activity analysis for those areas with ongoing conduct regulation if it is needed, which will in part depend on where the price-cap is being applied and the other issues that were discussed in Section 6; and
- for those areas where competition will be facilitated either through local access prices or cost information on existing operations then it is both feasible and important for Ofwat to consider the use of sub-company/activity analysis. This is because the focus will be on local networks or specific plant which are unlikely to have been the level where price controls were previously being set.

Again, it will be important to consider the cost benefit assessment to determine where sub-company/activity analysis should be undertaken, but when it is vital to facilitating competition then the likelihood is that the approach will need to be employed. As noted above, sub-company/activity analysis can be linked with panel data approaches and should be when the cost benefit assessment undertaken by Ofwat is positive – whether this is for transitional or ongoing facilitation of competition.

With respect to the available approaches to performance measurement, building on the comments in Section 6, the following comments can be made:

- econometric and frontier analysis approaches are likely to remain valid options; and
- index number approaches may become more applicable as competition increases the transparency of intermediate output prices which can be used in the estimation process.

It is also possible that Ofwat will be able to move, over time, to techniques like index numbers which have a lower regulatory burden for those areas where competition develops as the role of performance measurement will be as a regulatory check rather than an element of price determination or detailed information provision to facilitate competition. As such, there may be a short- to medium-term increase in Ofwat's use of detailed performance measurement approaches but then an ability to lighten the regulatory burden at least in some areas.

As such, Ofwat will need to consider the specific opportunities that arise as the industry structure beyond PR14 becomes clearer.

In addition, the possible changes in industry structure could have implications for cost interactions. These may exist between activities which become separated as part of the changing industry structure and imply the potential for biased cost assessment if not tested and controlled for. We therefore recommend that consistency checks are undertaken so that efficient estimates can be determined to encourage efficient competition.

7.4. Summary

Overall, it is likely that Ofwat will need to consider both panel and sub-company data beyond PR14. The precise requirements will depend on the final choice of industry structure and the way that competition develops. Ofwat should be able to use the framework developed in this report to assess those situations and determine exactly what is needed once there is greater clarity about the future of the industry and the associated market and industry structures after the publication of Defra's Water White Paper.

Establishing exactly what performance measurement is feasible and appropriate will depend on the way that the industry develops. We believe that:

- Ofwat will continue to need to set prices for the monopoly elements of the industry. As a minimum it will need to measure the performance of networks businesses for water and sewerage (as, for example, Ofgem continues to do with transmission and distribution); and
- Ofwat would need to conduct some form of regulation, probably price regulation for elements of the industry, such as franchise retail. Ofwat would also need to measure the performance of these elements. For these activities Ofwat should consider:
  - panel data when an ongoing need for conduct regulation exists and should carry out a cost benefit assessment to assess whether it is appropriate; and
  - sub-company and sub-activity analysis when there is a need for greater information to facilitate the development of competition markets.

While this analysis suggests that both panel and sub-company/activity analysis is feasible beyond PR14 whether they are actually used will depend on the cost benefit analysis, described in the following section.

8. COST-BENEFIT ASSESSMENT

Key themes and lessons

The fact that it is feasible to use panel and sub-company data does not mean that it is necessarily appropriate to do so. Rather it is necessary to consider the costs and benefits of the approaches in those situations.

Establishing quantitative costs and benefits for panel and sub-company models in possible future industry structures is not possible. Instead, we have used a qualitative approach to provide an overall assessment. In addition we need to understand the counterfactual against which we have measured any change. We have used the approach to performance measurement Ofwat took at PR09 and its existing data collection as the best available counterfactual, although the performance measurement and data collection requirements are unlikely to be exactly what would be used at future price reviews.

We believe that panel data has a positive cost benefit assessment in those cases where:

- there is little need to build historic datasets through retrospective data collection; and
- there is an ongoing need for price regulation/performance measurement which requires the use of both annual reporting and periodic forward looking price determinations.

Given these conditions, we believe that Ofwat should consider panel data for the "traditional" regulated elements, i.e. the network businesses and other essential facilities. There may be other cases where panel data is appropriate and Ofwat should assess those situations on a case-by-case basis, but as a minimum it should consider the use of panel data for the networks within the industry.

Our focus on using panel data for the segments of the industry with an ongoing need for price regulation creates a clear link across to sub-company data. In other words, to allow price setting for the natural monopoly elements the use of sub-company data will be important, such as treated water distribution. Additional data collection costs (to the regulated companies and Ofwat) are likely to arise through the collection of additional explanatory variables for output and input measures, and to control for heterogeneity, etc, where required. Initial additional costs for establishing this data collection may be high, while the additional ongoing costs will relatively low (compared to the existing requirements).

For sub-activity data we believe that the cost benefit assessment provides the following recommendations:

- Ofwat should consider using sub-activity data for improving cost transparency in contestable/potentially competitive segments of the value chain;
- Ofwat should consider using sub-activity data in setting cost-based access prices to facilitate competition. However, it is not clear that the benefits here will outweigh the costs, so Ofwat would need to consider on a case-by-case basis if the specific details/circumstances would be appropriate; and
- it is unlikely that conditions will be such that Ofwat should use sub-activity data for standard price determination efficiency analysis.

In those cases where circumstances for price setting, such as access charges, make the use of sub-company or sub-activity data appropriate we would recommend that Ofwat do this in conjunction with ongoing panel data usage.

The choice of modelling technique is dependent on Ofwat's objectives. We consider that OLS or DEA is generally a more transparent technique than SFA. However, there are benefits to an SFA approach such as explicitly attempting to take into account measurement error. We consider that, non-parametric index number techniques are less applicable at levels below the price cap, and Ofwat may prefer to use them as a consistency check at the price cap level.

Having considered the conditions in which it might be appropriate to use panel and sub-company data it is now possible to consider whether a cost benefit assessment supports their actual use. It should be noted that we do not believe it possible, or even meaningful, to try to quantify the costs and benefits associated with panel and sub-company data. The estimates would be dependent on the industry structure, form and level of regulation, approach to performance measurement etc. Rather, we have focused on providing a high-level qualitative assessment that seeks to establish an overall cost benefit assessment which allows us to see under what circumstances either of the approaches might deliver a positive assessment.

When undertaking a cost benefit assessment it is necessary to measure against a counterfactual. We have chosen the approach Ofwat took at PR09 as our counterfactual although it is not clear that this is necessarily the right one since it is not clear that this approach would, or could, be applied at future price reviews.

8.1. Panel Data

Costs

As has been discussed in this report, one of the key issues linked to either panel or sub-company data is that of data availability. When considering the costs and benefits of panel data the key cost that we need to assess is linked to data.

There are two ways in which data costs could occur:

- as a one-off cost through establishing a new consistent time series historic dataset at the outset so that sufficient minimum data exists for panel analysis to be undertaken; and
- on an ongoing basis as the dataset is updated each year.

Clearly in both cases the cost will depend in part on the level of aggregation that is sought. However, experience to date (e.g. UKWIR) has been that building historic datasets through retrospective data collection has proven difficult and costly. Ongoing data collection costs, while potentially high, would not necessarily show an incremental increase as the cost of additional data required for the panel data analysis is likely to be low or zero if the assessment is undertaken at the company level. In principle, the ongoing data collection costs may be less intrusive and costly than they currently are owing to the refinement and refocusing of regulation. If Ofwat were to make any change to the models and reporting requirements this would result in additional costs. A further point to consider however is that given the potential for additional (or different) performance measures required in future price controls, Ofwat may either need to collect additional data for the use in future controls or rely on shorter datasets once the performance measurement requirements have been determined.

There is also a risk that data will be collected which is then not needed for later determinations. As described in section 7, there are obvious areas where an ongoing need for conduct regulation will be needed and consequently these should be the focus for panel analysis. The risk of unnecessary data collection is then low.

Table 8.1 below summarises our views concerning panel data requirements and modelling issues.

Table 8.1: Implications of the modelling issues (including data requirements) for the assessment of panel data

Issue	Solutions	Impact on costs and benefits
Underlying economic theory		
Input costs	While potentially not a major concern, solutions include: Address through real price effects	This may impose some additional ongoing data collection costs, however we believe that these will not be significant (especially in relation to the impact on the allowed cost base) as high level price indices on different types of inputs could be used.[61]
	Regional price adjustments	Data is available from the ONS and Ofgem provides an example of the way in which this can be used. However, in some cases sufficient public data may not be available and Ofwat may need to source additional data.
Heterogeneity	Introduction of additional variables to capture aspects of heterogeneity	While in principle this could require additional information to be requested we do not think this is a significant issue for panel data per se because this would need to be considered even if cross-sectional data was used.
Opex-capex	Modelling of opex with quasi fixed capital or totex	Again, this is not a panel data issue per se. Additional data may be needed, especially relating to the capital stock, but that is probably true of any improved performance assessment model.
Other factors		
Technical change	Introduction of additional variables to capture aspects of technical change	While we do not believe that this is significant for the areas where panel data is likely to be used, it will necessitate further testing and the possible collection of additional information.
Cost interactions	Testing for cost interactions and possible use of additional models at aggregated levels	This is not an issue for panel data per se, but could place some additional periodic cost on Ofwat as more modelling would be needed than is currently the case (although any improvement in performance assessment is likely to incorporate some additional modelling).

[61] We note that Ofgem applied RPEs to some cost areas as part of DPCR5.

Scale	Use of appropriate model definition	Not a panel data issue per se because this would need to be considered even if cross-sectional data was used.
Length and breadth of panel	Retrospective data collection Consistency over time Pooled data analysis at PR14 Use of forward looking data	As discussed above, retrospective data collection could be costly. Ongoing data collection is likely to impose little or no additional cost compared to the existing approach. Data definitions can change over time. Ofwat would need to ensure that it understands the basis on which companies are collecting and reporting data to it. Explanatory variables may be updated over time and Ofwat may need to take this into account (if these variables are used in its modelling). An overlap year could be collected when definitions change. This would allow for an adjustment for structural breaks in the data. No additional cost but a reduction in likely benefits compared to a panel data specification. Ofwat could use data which companies submit in their business plans for other purposes, so there would be no additional cost. However, we note that there is little evidence on the use of forward looking data for benchmarking, and it would be up to Ofwat to consider the feasibility of it in their modelling.

Table 8.2 summarises our views, based on table 8.1, of the costs, both for the companies (collection and consistency) and Ofwat (consistency and analysis) of utilising panel data analysis. The table takes the existing level of regulated company data collection costs (i.e. under accounting separation guidelines) as a starting point and then considers the implications for costs of building an historic panel data database and the ongoing costs of data collection. The table also considers the cost implications of collecting panel data at the sub-company and sub-activity level which shows that there would be an increase in costs even if this was not a panel data issue. As only accounting data is currently collected under the accounting separation guidelines at the sub-company level, there would be an additional cost to collecting explanatory variables for use in cost modelling at the sub-company and sub-activity level. This cost is likely to be reasonably high given the need to establish reporting requirements and regulated companies' initial attempts to gather this data.

Given the data that is currently collected at the regulated company level, in our view an ongoing panel data model would not impose significantly greater cost than is currently faced by the sector. However, if sub-company and sub-activity data has to be collected on an ongoing basis the cost will be relatively higher given the greater volume of data need to be collected and processed. Costs at the sub-company/activity level are discussed further in Section 8.2 below.

Table 8.2: Assessment of the data costs for panel data

	Initial 'set-up' cost	Building an historic database for panel data	Ongoing panel data model
At regulated company level	Low (a number of costs and explanatory variables are already collected at this level)	Low cost (Ofwat has collected data at this level for a number of years)	Zero or low additional cost (data would be collected annually)
At sub-company level	Medium cost (cost of collecting data at a more granular level e.g. explanatory variables)	High cost (retrospectively building a panel data set would be costly)	Low/Medium cost (cost of collecting data at a more granular level)
At sub-activity level	High cost (cost of collecting data at a more granular level e.g. accounting data, explanatory variables)	Very high cost (retrospectively building a panel data set would be costly)	Medium/High cost (cost of collecting data at a more granular level)

We have noted previously in this report (Section 2.3 and Section 4.4) that Ofgem is exploring the use of forward looking data. We consider that this is an area that Ofwat could also explore as it may offer the potential to lengthen a panel dataset. However, this would need to be approached with caution as there is currently little evidence on the practical issues and limitations associated with using forward looking data for regulatory benchmarking purposes.

Benefits

The key benefit from panel data is the expected improvement in terms of stability and model specification that arises from the longer dataset. These models should be no worse than the models Ofwat used at PR09 and ought to be considerably better in terms both of the information provided and in their robustness and stability, if controls are put in place for technical change, input substitution, etc (where required).

While not a cost, a potential limitation on the benefits from panel data may be an application issue linked to the problems posed by different price input inflation across comparators, and the need for controls for inter-temporal variation in operating characteristics. We do not believe that this would fully off-set the benefits associated with panel data. Moreover, if panel data allowed more direct modelling of exogenous company specific factors this would in fact be a substantial improvement on the current approach where limited observations hinder this approach.

Our assessment of the benefits of panel data is provided in Table 8.3.

Table 8.3: Assessment of the benefits of Panel data

	PR14	Beyond PR14
Number of explanatory variables that can be included	✓	✓✓
Ability to account for technological change	✓	✓✓

Where: ✓ is positive and ✓✓ is strongly positive.

While we believe the relative benefits for PR14 are limited, in part because of the lack of historical data which can only be partly overcome through forward looking information, they will be a lot greater beyond PR14. However, as discussed in Section 3, if appropriate consistency can be established between company level data available before 2009/10 and Ofwat collects data after the new accounting guidelines came into place, (including data on explanatory variables) there would be greater potential for panel data analysis at the water and sewerage level at PR14.

Overall position

Panel data has a positive cost benefit assessment in those cases where:

- the need for building an historic dataset through additional initial data collection is limited or non-existent; and
- there is an ongoing need for price regulation/performance measurement which requires the use of annual reporting.

Given these conditions, we believe that panel data should be considered for the "traditional" regulated elements for PR14, i.e. the pipes businesses and other essential facilities, as: (i) for the foreseeable future, i.e. beyond PR14, the traditional regulated elements will require a performance measure (for price cap purposes), therefore Ofwat can plan its data requirements for its cost models; (ii) the definition of the historic data collected by Ofwat is relatively consistent with the network related business units (although we note that power is removed and modelled separately, this data is still collected and could be reapportioned to the network related business units); and (iii) adjusting for scale, heterogeneity, etc is relatively more straightforward. The build-up of panel data over time will be helpful in improving performance assessment.

There may be other cases where panel data is appropriate (e.g. water resources) and those situations should be assessed on a case-by-case basis but as a minimum there is a clear need to consider the use of panel data for the networks within the industry. For price controls beyond PR14, if regulatory commitment has been made to consistent data collection for a number of years (at least three for data to have 'settled') then panel data should be considered for all separate cost areas where performance measurement is considered necessary.

8.2. Sub-company and sub-activity

As discussed earlier in this report, sub-company and sub-activity data could be considered for the following possible uses:

- setting price controls at the sub-company level;
- estimating efficiency at sub-company and sub-activity levels; and
- supporting the development of competition (either through improved cost transparency of potentially competitive activities or by facilitating the setting of more cost reflective access charges).

The cost benefit assessment needs to take these aspects into account.

Costs

As with panel data, there are two possible types of data issues that need to be considered: set-up costs for a new dataset; and then ongoing collection costs. There are two aspects to the cost associated with sub-company/ sub-activity data – the volume of data and ensuring consistency across companies. As mentioned in the Section 8.1, the cost with respect to building historic datasets of sub-company and/or sub-activity data is greater than that of panel data at the company level due to the greater volume of data required. Although limited sub-company data and, to a lesser degree, sub-activity data is currently collected by Ofwat. Ofwat has discussed collecting more of this type of data with companies and even if only one or two previous years of data is needed, the costs associated with ensuring that consistent estimates are provided are viewed as prohibitive by the companies. However, the significance of the initial costs of collecting the volume of data will depend on their existing internal and external reporting requirements. What is likely to be more prohibitive is ensuring data consistency across companies where no existing regulator mandated collection occurs since companies are likely to have different data definitions and the cost of aligning these can be prohibitive.

It is also likely that the ongoing collection costs of sub-company/activity data will range from low to high depending on the level of granularity. Regulatory reporting requirements would become more intrusive and may result in requests for information that the companies do not naturally collect for themselves or where although the data is currently collected it is not on a uniform basis between companies and consequently costs will be incurred to ensure data consistency. Given the greater volume of data required there could be an increase in ongoing collection costs above those associated with the existing data collection process.

Our assessment of the data needs and other impacts of the economic and modelling issues associated with sub-company and sub-activity data is provided in table 8.4.

Table 8.4: Implications of the modelling issues for the assessment of sub-company and sub-activity data

Issue	Solutions	Sub-company: Impact on costs and benefits	Sub-activity: Impact on costs and benefits
Underlying economic theory			
Input costs	See panel data discussion		
Heterogeneity	Introduction of additional variables to capture aspects of heterogeneity.	While in principle this could require Ofwat to request additional information we do not think this is a significant issue for sub-company data as at this level heterogeneity may not be that significant and, where Ofwat determines that controls are needed, it may be relatively easy to control for.	This is likely to require significant additional data collection at the sub-activity level and this could impose significant costs. Even if some of the data is currently collected by the companies there would be costs for ensuring that consistent definitions are used across companies.
Opex-capex	Modelling of opex with quasi fixed capital or totex.	Again, this is not a sub-company data issue per se. Additional data may be needed, especially relating to the capital stock, but that is probably true of any improved performance assessment model. Ofwat may also need to consider whether companies' depreciation policies are consistent.	Again, this is not a sub-activity data issue per se. Additional data may be needed by Ofwat, especially relating to the capital stock and this would be at a level of granularity beyond what is currently available. As such, the information costs are likely to be higher for sub-activity analysis.
Other factors			
Technical change	Introduction of additional variables to capture aspects of technical change.	Not an issue specific to sub-company data per se.	Not an issue specific to sub-activity data per se.
Cost interactions	Testing for cost interactions and possible use of additional models at aggregated levels.	Potentially an important issue for sub-company data analysis and will require testing by running models at higher levels of aggregation and could reduce the benefits of using sub-company data.	Potentially an important issue for sub-activity data analysis and will require testing by running models at higher levels of aggregation and could reduce the benefits of using sub-activity data.
Scale	Use of appropriate model definition.	No additional data issues but may reduce the benefits of sub-company data analysis.	No additional data issues but may reduce the benefits of sub-activity data analysis.

As noted in Section 8.1, it is clear that retrospectively building a database of historic sub-company data would be expensive, if at all possible. Collecting sub-company/activity data on an ongoing basis would be relatively more expensive than Ofwat's existing approach which explicitly splits cost data by sub-company, but may not collect sufficient information on explanatory variables. Ofwat collects little information on sub-activities (aside from a few specific sub-activities e.g. sewerage treatment plant information). Consequently the benefits of a sub-company approach compared to the PR09 approach would need to be significant and a sub-activity approach even more so.

Our assessment of the data costs, based on the impacts set out in table 8.4 below.

Table 8.5: Assessment of the data costs for sub-company data

	Initial 'set up' cost	**Ongoing data model**
At sub-company level	Medium cost (cost of collecting data at a more granular level e.g. explanatory variables)	Low/medium cost (cost of collecting data at a more granular level)
At sub-activity level	High cost (cost of collecting data at a more granular level e.g. cost, explanatory variables)	Medium/High cost (cost of collecting data at a more granular level)

Benefits

The benefits of sub-company data for standard price determination are clear and relate to better cost transparency, more robust modelling, etc; but for sub-activity the benefits are far from clear. While Ofwat has used this approach to overcome issues with data observation limitations for sewerage, economic theory, as discussed in Section 3, questions whether these results are meaningful or unbiased, especially for sub-activity. While that does not mean the results cannot be used, they need to be used carefully and preferably in conjunction with other benchmarking that helps address these limitations.

With respect to competition the benefits are clearer. When seeking to provide greater cost transparency sub-company or sub-activity data, if focused at the level of activity where new entry is possible (say at the level of a treatment works) then the improved transparency and additional cost information should help encourage efficient new entry. Further, this increased transparency is likely to only be required for a transitional period as the area of contestability moves into one where actual competition exists. Of course, there could be an ongoing need for monitoring but that could probably be done at a less intrusive level of reporting (or would need to be offset against the expected benefits from competition).

If sub-company or sub-activity data is needed to facilitate more cost reflective access prices then again there are likely to be benefits, but we are unsure of how significant they are. From an economic theory perspective it is clear that getting the right local access prices will help facilitate competition, but from a regulatory perspective what is not clear is how much of a distortion is created by not having precise local access charges. This will, of course, depend in part on how competition develops in the sector.

Consider the energy industry. It is the case that some locational signals are sent through the market design while access prices at both the transmission and distribution level may have some local nature, but potentially not to a significant degree (especially in transmission in some parts of continental Europe). The degree to which a network is interconnected etc. affects that access issue to a point, but what is clear is that regulators are happy to minimise the locational aspects under certain circumstances. Further, there are examples from other sectors, such as rail in Great Britain, where the efficiency analysis is undertaken at a national level but then access tariffs are set on a local basis (but incorporating the national efficiency issues). We acknowledge, however, that there are far greater local aspects to the water industry costs than in other industries. Consequently, while the lessons from energy and rail are informative, it is plausible that more local access pricing may be required in the water industry.[62]

As such, we are unsure of the degree of incremental benefit achieved by using sub-activity analysis for local access prices. Of course, if the price-caps are being set at the local level, then this should be the basis for some of the efficiency analysis (subject to the usual provisos about cost interactions etc). As such, only once it is clear what forms of competition will develop can Ofwat undertake a real assessment of whether sub-activity data for "local" access prices is required. It is, however, clearly an issue that Ofwat needs to keep under review and may prove to be an important use of sub-activity data.

Our assessment of the possible benefits of sub-company and sub-activity analysis is provided below.

Table 8.4: Assessment of the benefits of sub-company and sub-activity data

Level	**Purpose**	**PR14**	**Beyond PR14**
Sub-activity	Price setting	-/✓	-/✓
	Facilitating competition	-/✓	✓✓
Sub-company	Price setting	-/✓	-/✓
	Facilitating competition	✓	✓✓

Where: ✓ is positive, ✓✓ is strongly positive, and '-' is neutral.

In part our assessment of the price control beyond PR14 is predicated on the fact that it should be clearer as to how competition is developing in the sector and so where more information would be appropriate and whether local access prices are needed.

Overall position

Given the three possible uses of sub-company and sub-activity data we believe that the cost benefit assessment provides the following recommendations:

- Ofwat should use sub-company data to set price controls for network/essential facility services in the long-term (alongside panel data);

[62] See for example the discussion from Bolt (2010). Available from the Ofwat website.

- Ofwat should use sub-company and sub-activity data for improving cost transparency in contestable/potentially competitive segments of the value chain probably as a transitional action;
- Ofwat should consider sub-activity data when considering setting access prices to facilitate competition but it is not clear that the benefits here will outweigh the costs, so a case-by-case consideration of the specific details/circumstances would be appropriate or the assessment should be kept under review as the way in which competition develops becomes clearer and the implications of not having efficient local access prices can be determined more precisely; and
- it is unlikely that conditions will be such that sub-activity data should be used for standard price determination efficiency analysis.

In those cases where circumstances for price setting, such as access charges, make the use of sub-company data appropriate we would expect this to be done in conjunction with ongoing panel data usage.

8.3. Cost modelling technique

We have discussed the particular key attributes of three different cost modelling techniques in Section 3.4 (and Annex A.7), namely, econometric, DEA and non-parametric index numbers. We also noted in Section 4 that regulators have tended to use less sophisticated techniques.

While we do not presume to recommend an approach to Ofwat, we consider that econometric average response techniques are generally well understood by regulators and regulated companies and are much more transparent than, for example SFA. However, we note that ORR do use an SFA approach and recent academic literature shows how this technique can be applied to water and sewerage (for examples, see Annex A.7). DEA and non-parametric index numbers are also relatively transparent and reproducible, however given non-parametric index numbers relative inflexibility to deal with different operating environments and it requirements of output price indices (or proxies) we consider that its use may be more restricted to consistency checking at the price cap level.

9. CONCLUSIONS AND OBSERVATIONS

Key themes and lessons

While this report has focused on establishing an approach to evaluating when panel and sub-company models are appropriate, through a decision tree, we can draw some clear conclusions and make some useful observations.

Traditional network access regulation for pipes and other essential facilities will, for the foreseeable future, require price regulation and some form of performance measurement. For these aspects of the value chain we believe panel data modelling benefits outweigh the costs, and Ofwat should strongly consider panel data use in future price controls. This is particularly the case as the available data builds up beyond PR14.

For those areas where Ofwat wants to encourage and support/facilitate competition, sub-company approaches are likely to be appropriate. It is only in those areas where greater cost transparency and possibly local access prices are needed that the benefits of sub-company approaches will outweigh the significant costs of collecting the additional data.

Given the changes that may potentially occur in the sector and the way regulation is developing, it is important that Ofwat introduces new consistency checks to ensure that it can address any potential biases in estimates owing to the existence of cost interactions.

There are a range of possible approaches that Ofwat should consider developing. These will both help establish the consistency checks as well as allowing the development of other approaches to performance measurement.

9.1. "Traditional" network access regulation – panel data

Regulatory commitment (and ensuring data comparability and quality over time) has provided regulators in the UK and internationally with relatively reliable and robust panel data sets that have increasingly been used for performance measurement – particularly for networks. While there is scope to use panel data for other areas of the water and sewerage sector, we consider that the clearest case is for the pipes businesses where, for the foreseeable future, some form of performance measure will be needed. This should allow Ofwat to commit to consistent data collection in this area. Ofwat's data collection has provided it with a relatively stable high level dataset, i.e. total *water* opex and capex, and total *sewerage* opex and capex. There is the potential for such data to be consistently employed for the period before and after the accounting separation guidelines were introduced. Over time, data collected under Ofwat's accounting separation guidelines will also yield a stable dataset available for use at PR14, which, provided that data consistency is maintained, should expand into a longer panel data set beyond PR14. We, therefore, consider that the use of panel data in Ofwat's cost modelling offers the potential to improve on the approach used for PR09, in relation to the robustness of the results and the ability to control for a greater range of factors (e.g. input substitution, technical change).

However, given the scope of the cost allocation requirements there will likely be a need for Ofwat to carry out cost allocation consistency checks. We consider the following factors are important considerations when deciding on the use of panel data given that traditional network access regulation will almost certainly require, some form of efficient cost determination on an ongoing basis, beyond PR14:

- relatively low number of comparators (cross-sections) suggests the need for panel data as this increases the confidence in the results (i.e. more robust model outputs); and
- as panel data combines data across multiple years, models will require controls for changes in operating characteristics, input prices, and allowances for technical change.

If regulatory commitment is made early enough to the collection of data at the vertical segment level (i.e. as in accounting separation) then at PR14 (and beyond) there should be a sufficient number of observations to produce robust sub-company cost models. However, whether these are needed will depend in part on the decisions about the design of the regulatory regime.

Given the above, we recommend that Ofwat strongly consider the use of panel data for modelling traditional network access regulation for future price controls, with appropriate checks for data consistency over time and across comparators.

9.2. Facilitating competition – sub-company

Undertaking comparative benchmarking at levels below the price cap is theoretically feasible for discrete operating units and licences, at least if a sufficient number of stable and accurate observations is available. However, given the additional data requirements and the difficulties of creating models at a level below the price cap, we recommend that Ofwat's use of comparative benchmarking below the price cap level should be focused on the need to facilitate competition and for determining appropriate access prices. As discussed in Section 7, we believe this will be increasingly important beyond PR14.

Because there is considerable variation in network and treatment costs attributable to local operating characteristics, sub-company data and comparative benchmarking is potentially required. Panel data can be useful for sub-company benchmarking; however, the data for this needs to be both stable and accurate.

9.3. Ensuring consistency in cost estimation and allocation

As we have discussed in Section 3 and illustrated in Section 6, if analysis is done at a cost level below that of the aggregated companies' costs then issues around cost allocation, and economies of scale and scope need to be considered. Therefore, if analysis is done below the company level the following checks should be undertaken:

- *Consistency checks* for aggregation bias, impact of vertical and horizontal integration economies and other cost interactions.
- *Transitional performance assessment* required to ensure that any transition from aggregated price caps to disaggregated price caps is done with consistent cost assessment measures.

9.4. Methodologies that facilitate reduced regulatory burden and consumer side allocative efficiency

The academic literature (see Annex A), CC decisions, and the experience of other regulators suggest that possible alternative methodologies, which may lower regulatory burden but still meet regulatory objectives, should be considered. This is also consistent with some of the expected recommendations of the Gray report for Defra with respect to lightening the regulatory burden, something Ofwat has separately committed to seek to deliver.

The application of panel data would allow Ofwat to assess performance at a level of company aggregation more closely linked to the level of price caps than the approach currently applied.

This strongly suggests that the use of panel data has the potential of simplifying the regulatory process by aligning the unit of performance measurement with the unit where regulation is applied. Thus, appropriate panel data estimation of highly aggregated opex or total regulated cost models, that are estimated with due respect for economic theory (i.e. taking account of flexible production models, heterogeneity, economies of scale, cost allocation issues, etc), has the potential to reduce regulatory burden. Moreover, the data requirements for such modelling are likely to be less intensive than Ofwat's current approach. Thus, if such modelling were coupled with assessment of key performance indicators on key outputs/outcomes it could potentially provide an effective approach to yardstick regulate companies. This could be achieved by monitoring relative firm performance, determining high level assessments of past and potential X factor performance, and imposing less intrusive regulation subject to the condition that firms meet their X factor and KPI targets.

We finally speculate with regard to the potential for Ofwat to employ panel data techniques in order to jointly assess the relationship between a firm's cost performance, profitability, and ensuring that consumer prices are cost reflective in the long term. Particularly if we see the introduction of more competition and/or alternative regulatory models, Ofwat might consider the further development and adaptation of firm level profit decomposition techniques such as that illustrated in De Witte & Saal (2010). This paper illustrates a panel performance measurement technique that not only provides a decomposition of the sources of firm productivity growth (including efficiency and technical change), but also assesses its impact on firm profitability and hence consumer oriented allocative efficiency. While this paper is implemented with a DEA based cost assessment, there are also academic examples of such profit decomposition techniques which employ index number and econometric estimation, and which could potentially be developed to satisfy such regulatory requirements.

ANNEX A: WATER AND SEWERAGE COST MODELLING: AN ACADEMIC ASSESSMENT

This annex provides a more detailed academic review of relevant issues related to the economics of cost modelling in the water and sewage industry and provides the background for the summary provided in Section 3 of the main of the report.

A.1. Motivation for the academic assessment

There is a substantial difference between the theory and practice of economic cost modelling in the academic context as opposed to the regulatory context, where cost modelling has developed as a necessary, pragmatic and effective tool for assessing the relative performance (and therefore efficient cost assessment) of regulated water and sewerage companies. Moreover, many of these differences are the result of differences in objectives between regulators and academics. Firstly, the primary objective for regulatory cost assessment in the water industry has been relative efficiency assessment (with a balance between accuracy and discretion), while the academic literature has focussed more on the impact of privatisation, and changes in regulation and industry structure on productivity growth and firm performance. Secondly, given its regulatory requirements, Ofwat has historically assessed a company's overall allowed regulatory costs with a building blocks approach that aggregated separate assessments of required capital expenditure (capex) required operating expenditure (opex) and an appropriate return on an allowed regulatory capital value (RCV). This approach reflects, not only the quasi-fixed nature of capital in most infrastructure industries, but also a necessary regulatory approach given the post-privatisation regulatory settlement of RCV values. In contrast, many academics model total economic costs or variable costs (opex), with modelling approaches that attempt to capture the potential for input substitution between capital and other inputs, which economic cost modelling theory emphasizes. In a final difference, Ofwat's regulatory cost assessment models have almost exclusively employed cross-sectional approaches, while academic cost modelling has increasingly relied on panel data assessment.

Given these and other substantial differences between the objectives and requirements of regulatory and academic cost modelling, there are clearly appropriate and necessary differences between the types of models employed. Therefore this section does not set out to advocate the wholesale adoption of an academic approach for regulatory cost assessment. However, we nonetheless believe that there are useful insights and considerations that can be drawn from the economic theory and academic practice of cost modelling,[63] and this is particularly the case given Ofwat's desire to consider the potential application of panel data models in regulatory cost assessment.

This annex will therefore briefly consider the economic theory of cost modelling, as well as empirical examples, which given the focus of this report on water industry cost modelling, are largely drawn from the academic literature on water industry costs in England and Wales. This will allow us to highlight several generic modelling issues that may influence model accuracy in

[63]See for example, Hackman (2008) for an excellent recent treatment of economic production and cost theory, as well as efficiency, productivity, and performance measurement.

both the academic and regulatory context, and particularly issues that may have an impact on the use of panel data in future regulatory assessment   However, we should also emphasize that this annex does not strive to provide an exhaustive discussion of academic cost modelling nor the empirical literature on water and sewerage cost modelling. Instead, the purpose is to briefly review the academic literature, with the intention of highlighting those issues that are most likely to inform and influence the potential application of panel and sub-company modelling, as well as carrying them out with accounting separation data, to regulatory cost assessment in England and Wales.

A.2. The basic economic theory of cost modelling

The basic economic theory of costs rests on the assumption that firms aim to minimize the long run total cost of production given the production technology available to them, the outputs they produce, and the input prices they face.[64] It also highlights the need for fully flexible modelling approaches able to allow for a fuller specification of the underlying relationship between costs, outputs, input prices, and the scale of operations. As a result, economic cost modelling generally requires the specification of costs as a function of outputs and input prices, and the use of an empirical model that is sufficiently flexible to reflect how alternative methods of production influence a firm's costs. Fundamentally, such flexible models allow for the cost implications of important economic relationships to be captured in the empirical modelling. To illustrate this, we focus on the economic concepts of input substitution and economies of scale, which are particularly relevant in the English and Welsh water industry.

A.2.1. Input substitutability, input prices and panel regulatory cost assessment

Input substitutability implies that if firms have the potential to substitute one type of input for another (such as capex for opex), they will strive to adjust their input usage so as to reduce their total cost of production by using more of the relatively cheaper inputs. This also implies that firms which face different input prices may have differences in costs. Moreover, such input substitution is highly relevant in the water industry given frequent discussion of capital opex tradeoffs in the regulatory context. These effects have also been demonstrated in the academic literature analysing English and Welsh water companies with panel data. Thus, for example, Saal & Parker (2000) provides evidence of considerable gains in WaSC cost productivity attributable to the substitution of other inputs with capital, while Bottasso & Conti (2009) demonstrates that opex costs are influenced by capital stocks. Similarly, Maziotis, Saal, & Thanassoulis (2011) suggests that between 1994 and 2009, such input substitution has had a considerable impact on reducing WaSC costs.

From a cost assessment perspective, the potential for input substitution and the incentive for managers to minimise costs given the input prices they actually face, suggests the need for careful controls of input price variation between firms particularly when considerable cross-sectional variation in relative input prices exists. Moreover, the need for such input price controls becomes even more important in a panel context, as argued by Coelli, et al. (2003). This is because it becomes less plausible to assume that input prices will not change across time, and

[64] See, for example, Chambers (1988).

that relative input prices, for either the inputs used by a single firm, or those reflecting price differences across firms will stay constant across time.[65] Thus, any movement to the use of panel data in regulatory cost assessment will require substantially greater consideration of the impact of input prices on assessed costs. Failure to do so may result in biased estimates in firm costs, biased estimates in cost trends, and a substantial weakening of the potential to improve regulatory cost assessment through the use of panel techniques. Stated differently, allowing for input substitutability and input prices is desirable in order to allow for allocative efficiency as well as technical efficiency in determining relative firm cost. E.g. if firms face different input prices their rational cost minimising choice of inputs will be different and we should allow for this in an appropriately specified cost function model. The main body of the report therefore considers further the potential for this in the regulatory context, taking into account related issues such as data availability, any resulting regulatory burden and the potential cost relative to improvements in accuracy this would entail.

A.2.2. Model flexibility and the economic relationship between scale and costs

Sufficient model flexibility is also required in order to properly assess the relationship between a firm's output scale and its costs. Thus, for example, it is common for regulators to employ models of logged costs with logged outputs, and it is often assumed that such models allow for the presence of variable returns to scale. However, this is not in fact the case. Such log-log models allow for increasing, decreasing, or constant returns to scale, but impose the unrealistic restriction that all firms, regardless of their scale of operations, operate with **precisely** the same returns to scale. Only fully flexible forms such as the translog cost function, allow for truly variable returns to scale, and are therefore able to result in estimates reflecting the more tenable likelihood that smaller firms would benefit more than larger firms from increases in their scale of operations. Moreover, as both Coelli, et al. (2003) and Greene (2008) argue, even if efficiency determination is the primary purpose of cost modelling and scale economies are of only tangential interest, use of an appropriate flexible cost function able to adequately allow for potential variation in scale effects is essential as overly restrictive forms may result in distorted efficiency measures.

We believe that this issue of potentially distorted performance measurement is worth Ofwat's further consideration, given that recent academic literature reviews (Abbott & Cohen, 2009 and Saal, et al., 2011) suggests that while there are economies of scale for small water and sewerage firms, these economies of scale are exhausted at relatively modest firm sizes. Moreover, work commissioned and published by Ofwat, also suggested the presence of variable returns in the English & Welsh water industry, with the smallest firms characterized by increasing returns to scale and larger firms in the industry characterised with decreasing returns to scale.[66]

The cost assessment implications of allowing for variable returns to scale are considerable. One could draw on the data envelopment analysis (DEA) literature to make an argument that regulators should assess performance based on the costs of a firm operating at the optimal scale, and thereby penalise firms that are below and above this scale. However, this argument ignores

[65] For example, the relative price of labour to capital tends to increase over time, in tandem with a steady substitution of labour with capital in most production processes where this is feasible.
[66] Stone & Webster (2004a).

the fact that regulated utilities are not fully in control of their scale and must serve their licensed operating area and its population. Thus, utilities should be assessed with a methodology that can consistently capture the cost implications of their scale of operations. Moreover, as current regulatory cost assessment models do not allow for fully flexible returns to scale, consideration of the academic literature suggests that they may result in biased estimates of the costs associated with various scale levels, and may thereby create perverse incentives for firms to merge. Moreover, we would emphasise that panel data could provide Ofwat the necessary degrees of freedom required to more closely model the relationship between scale and firm costs.

A.3. Quasi-fixed capital stocks, the appropriateness of total cost modelling, and the modelling of opex and capex for regulatory cost assessment

While the concept of input substitutability and the resulting emphasis on allowing for input prices, as discussed in Section A.2.1., is an important and relevant concept for water industry cost modelling, and particularly so with the potential use of panel data, consideration of the past academic literature suggests that this approach is not always appropriate, as in the short run, some inputs may not be fully under managerial control. In particular, given the reality of infrastructure industries, we must consider how to properly model those variable costs and inputs under managerial control (most opex expenditures) when it is infeasible for managers to quickly adjust capital stocks (perhaps proxied by Modern Equivalent Asset (MEA) capital stocks) to their long-run cost minimising level. Stated differently, the magnitude and lumpiness of capital investments coupled with the long lived nature of capital assets in the water industry suggests that while capital stocks are always in the process of being adjusted towards some desired long run optimal state, managers are unlikely to reach this optimal state in a single year or even over many years.

Nevertheless, this process of incrementally adjusting the quasi-fixed capital stock has resulted in an ongoing increase in the quantity and quality of capital available to managers, which implies that, particularly in a panel context, these changes over time will have an impact on a firm's opex. Thus, for example Stone & Webster (2004b) found evidence that changes over time in both water and sewerage capital stocks had a considerable impact on opex requirements, and hence opex productivity growth rates.[67]

The theory and practice of economic cost modelling suggests that the appropriate economic model for such circumstances is not a long-run total cost model where all inputs are assumed to be fully variable. Instead, a model of variable costs (opex) which nonetheless allows for the influence of such "quasi-fixed" capital stocks on variable costs is more appropriate.[68] Moreover, this approach, which models variable costs as a function of outputs, variable input prices, and quasi-fixed capital stocks provides a theoretically appropriate approach to opex modelling while also allowing for the significant impact of capital opex tradeoffs. We also emphasise that particularly in a panel context, appropriate cost assessment of opex requires the application of such an approach, if we wish to accurately model cross-sectional and intertemporal variation in efficient opex levels.

---

[67] A similar quasi-fixed capital approach has recently been employed with WaSC data by Bottasso & Conti (2009a).

[68] See for example Gort & Sung (1999).

Given these arguments, what are the implications for regulatory cost assessment in general, and in the panel context in particular? Firstly, total cost assessment has a role, particularly with long panel data models of 10 years or more such as those commonly used in academic assessments, and can help identify overall trends in the full economic costs of activities. This is the case as the assumption of full flexibility in capital allocations is more reasonable in this context. However, while such total cost models can provide a useful cross check of firm performance, the quasi-fixed nature of capital stocks as well as the current five year regulatory cycle, suggests they should not be the primary tool of regulatory cost assessment in the water industry.

Instead, consideration of the academic concept of quasi-fixed variable cost modelling, lends support to the continued application by Ofwat of its current building blocks approach, in which overall regulated costs are obtained with separate assessments of opex and capex are coupled with an allowed rate of return on the RCV. This approach is appropriate, as it embodies the quasi-fixed nature of the capital stock, while allowing for the appropriate regulatory assessment of the efficiency of both operating costs and efficiency in capital expenditures.

However, while consideration of the academic literature suggests the continued assessment of both opex and capex, it also strongly suggests the need for better controlling for capital stocks as a determinant of opex, as changes in capital stocks still have a potentially significant impact on opex expenditures in a quasi-fixed capital setting. This will become more important in the panel context, where failure to allow for the impact of changes in capital stocks over time may result in biased estimates of required opex, biased estimates in opex trends, and a substantial weakening of the potential to improve regulatory cost assessment through the use of panel techniques. Moreover, by allowing for the impact of capital stocks within opex cost assessment, the regulatory cost assessment regime could potentially better capture the impact of capital opex tradeoffs on overall assessed costs.

A.4. The separability of multiple output production, sub-company modelling and cost interactions

As this report considers the potential application of sub-company data assessment and panel data techniques as well as the potential for employing such approaches with accounting separation data, we now consider relevant economic theory and empirical water industry work which relates to these issues. We therefore first consider the economic theory of multiple output cost modelling to provide insights with regard to cost assessment with accounting separation data, and then consider the implications of sub-company assessment from an economic perspective.

A.4.1. The separability of multiple output production, vertical cost interactions and their implications for accounting separation

The production of both water and sewerage services involves a complex process that is not easily characterised by a single output. Thus, for example, the academic literature provides recent papers in which the aggregate provision of water services by a WoC is modelled as the joint production of water volumes, customer connections, and water transportation.[69] Similarly, Ofwat's past modelling approaches, and particularly the separation of activities in the accounting separation guidelines that came into effect in 2009-2010, provide alternative characterisations of the multiple outputs produced along the water and sewage supply chains. Moreover, consideration of even a single output component of the accounting separation chain reveals that firm activities such as "water treatment" or "sewerage treatment" are in fact the horizontal summation of the multiple outputs produced within the firm's multiple treatment facilities, each with different output characteristics.

While full consideration of its theoretical detail is beyond the scope of this review, the seminal work of Baumol, Panzar, & Willig (1982) suggests that careful consideration of potential cost interactions between these multiple outputs is necessary if we wish to properly assess costs in a multiple output context.[70] Similarly, Chambers (1988) and Berndt & Christensen (1973) provide theoretical contributions with regard to the nature of multiple output production and the conditions under which the separate estimation of costs for outputs is appropriate. However, if the costs of two activities are economically nonseparable, unbiased cost assessments of one activity cannot be obtained without consideration of the other activity's influence on the first activity's costs.

In a paper cited in a wide variety of subsequent industry studies, Evans & Heckman (1984) demonstrated the empirical implementation of one possible test for the separability of outputs in a translog cost function, and rejected the hypothesis that local and long distance costs in the US telecommunications industry were separable. Focussing more closely on the water and sewage industry, Saal & Parker's (2000) direct application of Evan & Heckman's approach suggests that water and sewage activities are nonseperable. Moreover, the primary contribution of both Saal & Parker (2006) and Bottasso, et al. (2011) is to suggest a strong potential bias in the assessment of the efficiency of one activity (water only), if we ignore what appear to be substantial negative cost interactions between the water and sewage activities of WaSCs. Thus, consideration of previous academic literature on the separability of cost functions with multiple outputs suggests that there is the potential for biased cost assessment if such cost interactions exist.

We finally note that a substantial body of academic literature has considered not only whether cost interactions between multiple outputs exist (nonseparability) but whether such vertical cost interactions imply higher (or possibly lower) costs when outputs are produced in a separated as opposed to an integrated firm. A recent literature review by Saal, et al. (2011) summarizes the findings of approximately 15 published academic papers that have directly considered this issue for the water industry. This past literature, as well as the recent contribution by Saal, Arocena, & Maziotis (2010), suggests that such vertical integration relationships may be present between many components of the water and sewage supply chain.

---

[69] See the recent literature review of Saal, et al. (2011) for an international review of a substantial number of such studies. Two recent examples of studies based on English and Welsh data are Bottasso and Conti (2009b, and Saal, D S, Arocena, P & A Maziotis (2010).

[70] As the Baumol et al. (1982) report focussed on cost assessment, we largely leave aside this work's central contribution with regard to evaluating how the cost relationships between multiple outputs influence the relative cost of producing in traditional natural monopoly structures as opposed to industry structures designed to facilitate contestability with partial or complete vertically and horizontally separated structures.

In summary, our brief academic review of multiple output cost modelling suggests that if vertical cost interactions exist there is potential for an overall cost assessment to be biased unless the partial assessments properly account for any cost interactions between other activities which influence the costs of the activities under assessment. This suggests the potential need for a regulatory cost assessment process of:

- identifying and testing for the presence of any relevant cost interactions between activities;
- employing appropriate controls to account for such cost interactions;
- employing cost assessment at a level where such interactions are internalized;[71]
- using higher level assessments as consistency check of disaggregated assessment; and/or
- employing appropriate regulatory discretion to mitigate any potential biases in assessment.

It needs also to be recognised that these vertical interactions are dependent on the industry and market structure and are likely to change, possibly substantially, were the structure of the industry to change e.g. to support greater competition, upstream and/or downstream. This implies testing carefully for structural breaks in the relationships discussed above.

A.4.2. Consideration of sub-company modelling and potential horizontal aggregation bias

The use of sub-company data with disaggregation along a horizontal dimension, such as treatment plant data, is subject to a different potential bias in performance assessment to that discussed immediately above. If, as the existence of water and sewage companies with multiple plants suggests, horizontal integration economies are obtained by aggregating treatment plants, water supply zones, and other sub-company operations (at least to some optimal level of horizontal scale), assessment at plant level may not properly allow for these benefits of horizontal integration and, if so, this would result in upwardly biased estimates of aggregate costs for larger companies if horizontal integration economies are present. In contrast, if diseconomies of horizontal integration are present, larger firms might have downwardly biased estimates of aggregate costs. In other words, important cost interactions and implications for economies of scale/integration may not be properly accounted for by simply aggregating the cost assessments for sub-company data, unless they are properly controlled for in the aggregation process. As there is evidence to suggest that both economies and diseconomies of horizontal integration are currently extant in the water industry (Stone & Webster, 2004a; Bottaso & Conti, 2009a) sub-control measurement may simultaneously provide respectively downwardly and upwardly biased estimates of the efficiency of firms that operate with increasing or decreasing returns to scale. This therefore, suggests that the employment of sub-company data does not in practice provide

[71] We note that this particular approach was recently advocated by Oxera for the modelling of water opex costs by Oxera in its submission to the Competition Commission as part of the Bristol Water Inquiry. Thus, Oxera argued that the aggregate assessment of total opex as a function of the determinants the variable weren't the same as in the current water models (some were, but some weren't;) used in the four separate opex models employed by Ofwat in the 2009 price review, would internalize any cost interactions between these activities.

a straightforward means by which to assess the performance of firms, if for example, mergers substantially reduce the number of firm level observations available to Ofwat.

Our review of the previous academic literature suggests little previous guidance with regard to the potential magnitude of such possible horizontal aggregation bias, nor methods to test for and subsequently correct for any potential bias in company level assessment that might result from sub-company assessment. We therefore suggest that Ofwat consider further its experience with sewage treatment modelling, and whether there is a need to test for and control for horizontal aggregation bias, if sub-company modelling were to be more widely adopted.

A.4.3. Summary on vertical and horizontal cost interactions

Stated most simply, the presence of vertical or horizontal cost interactions suggests that assessment at the aggregated level, rather than with vertically and/or horizontally disaggregated data is most suitable as it will ensure that such cost interactions are controlled for. Thus, under any form of vertical separation (accounting, functional, legal, etc.), our academic review suggests that policy makers need to ensure that any potential bias in cost assessment is considered, tested for, and controlled for. The latter point is therefore an important issue that should be considered in the development of new cost assessment tools by Ofwat. Failure to do so might, for example, result in biased network access pricing, which, if prices were set above efficient levels may forestall competition, while in contrast if access prices are set below efficient levels, they might incentivise inefficient market entry.

A.5. Heterogeneity

Our consideration of economic cost modelling to this point, has assumed that despite variations in the scale and scope of activities carried out by firms, and the potential for cross-sectional and intertemporal variation in input prices, the underlying production technology and operating characteristics faced by all firms are the same. However, this assumption is particularly untenable in the water and sewerage industry where differences in topography, settlement patterns, water source type and location, required water and sewage treatment technologies, sewerage discharge location, variable discharge consents, and a great variety of other operating characteristics result in substantial and legitimate differences in firm costs. As a result, while Ofwat's current cost assessment practice does not fully control for such heterogeneity in the cost assessment models, it does allow for special factors, which the recent Bristol Water Competition Commission (CC) case revealed amounted to up to 14% of modelled water costs at PR09. Thus, cross-sectional heterogeneity is already an extant issue that impacts on regulatory cost modelling

As discussed briefly in Section A.2. above, heterogeneity will exist in both the cross-sectional and inter-temporal dimension. Operating characteristics such as customer demographics, water source reliability and quality, mandated water and sewerage treatment levels, operational quality of networks (e.g. leakage, pressure), etc. will vary across time, and such variation will increase with the length of the panel data set. Thus, any movement to panel data models will require the development of appropriate modelling approaches to properly control for such variation, and its impact on efficient costs at the level of assessment. Coelli, et al. (1999) provide a simple example of not only how such heterogeneity influences estimated efficiency, but also demonstrates how the modelling approach employed to control for it influences estimated inefficiency. Moreover, Green (2008) details further approaches to modelling heterogeneity with stochastic frontier analysis (SFA) and also discusses the effect of mishandling heterogeneity on inefficiency measurement.

Classical panel econometric models such as the random and fixed effects specifications allow for a firm specific error term in addition to the standard white noise regression residual, and therefore provide a form of control for firm specific heterogeneity. However, there is the risk that this heterogeneity includes persistent inefficiency as well as differences in operating characteristics. Thus, recent developments in panel techniques such as Greene's (2005) true random fixed effects model attempt to allow for inefficiency as well as firm specific heterogeneity. Application of this approach by Saal, Parker, & Weyman-Jones' (2007) to a WaSC database, reveals that even after allowing for a fixed group specific heterogeneity term, firm specific operating characteristics still significantly influence input requirements in the water sector. Thus, as appropriate controls for cross-sectional and intertemporal variation in operating characteristics will require the inclusion of further explanatory variables in assessed models, the additional degrees of freedom provided by panel data, will not increase perfectly in line with the length of the panel data set. This again highlights that the necessary modelling considerations associated with panel data imply that while panel models have the potential to improve cost assessment, they should not be naively assumed to simply operate as a degrees of freedom multiplier of a cross-sectional specification.

We note that a significant advantage of including appropriate heterogeneity controls directly in the cost assessment model is the ability to allow the model to determine whether a "special factor" is in fact a significant determinant of cost variation between firms. Thus, pursuing such an approach should allow Ofwat to provide a more accurate assessment of both the appropriateness of special factors as well as the direct quantification of their impact on estimated costs. However, properly implementing this approach will require retention of sufficient cross-sectional observations to facilitate it.

We finally consider the significance of heterogeneity in operating characteristics in relation to the costs and benefits associated with sub-company modelling. Focussing first on the benefits, we highlight that given that substantial heterogeneity exists at sub-company level within firms, sub-company assessment would allow more appropriate cost assessment of two essential costs which are required to regulate for competition. Thus, as suggested in Ofwat's January 2011 assessment of hypothetical upstream water markets,[72] efficient competition will require access pricing at the sub-company level, and location specific information on an incumbent's marginal costs in potentially contestable markets. Sub-company modelling, would in principle provide useful information to meet both of these needs. However, such modelling would not only require appropriate aggregation and consideration of horizontal and vertical cost interactions, as discussed above, it would, for completeness, also require collection of a substantially larger regulatory database with appropriate information on locally specific costs and operating characteristics. Thus, while there are clear benefits associated with this approach, the existence of even greater heterogeneity at the sub-company level than at the firm level implies that the costs associated with sub-company modelling will be substantial. Thus, from an academic point of view policy makers will need to evaluate whether the necessary estimates required to facilitate

[72] Ofwat (2011b).

efficient market entry have sufficient benefit to overcome the substantial costs associated with their unbiased estimation.

A.6. Allowing for technical change in panel models (cost function parameter change)

In an approachable text aimed at regulatory practitioners, Coelli, et al. (2003), argue for the application of sophisticated panel techniques, capable of capturing the cost reducing impact of technical change on assessed costs. Moreover, it is also the case that such approaches generally allow for technical change to be neutral, input augmenting, and/or output augmenting. This, consideration is important as, for example, it would explicitly allow for the reasonable possibility that as technology changes over time, the underlying cost function shifts and becomes more capex and less opex intensive, an observation which is borne out by the vast majority of panel based academic studies of water industry costs. This further relates to our above discussion of opex modelling with controls for quasi-fixed capital stocks, e.g. if capital deepening technical change is occurring over time, efficient opex levels will become lower over time in a panel database, all other things being equal. Thus, such capital deepening technical change should be allowed for in a panel based opex cost model, otherwise we may produce upwardly biased estimates of future efficient opex requirements.

Academic approaches also suggest that if panel models are to be used in any form of regulatory cost assessment, careful consideration of the nature of technical change, and its relationship to other modelling characteristics is necessary. Therefore, if no evidence of technical change is found, simple pooling of a dataset is appropriate. In contrast, if evidence of neutral technical change is found, which effectively implies that the relationship between costs and all inputs and outputs is changed in equal proportion, the simple inclusion of time dummies or a time trend may be appropriate. However, if technical change effects inputs differently, by for example favouring the use of more capital, or effects the costs of producing outputs differently, input and/or output biased technical change should be allowed for in the specification. Thus, while we cannot provide any ironclad rules with regard to the necessary controls for technical change in a regulatory panel based specification, we do believe that careful consideration of the types of technical change that might have influenced an assessed activity is crucial.

We finally note that academic literature suggests that technical change cannot be viewed in isolation from other modelling characteristics. Thus, for example, if technical change is capital biased, e.g. greater substitution of operating costs with capital over time, it may not be feasible to detect such technical change, unless a model allows for capex-opex substitution. We therefore further explore the implications of this potential pitfall in our next section, by discussing it in relation to econometric approaches to performance measurement.

A.7. Empirical approaches to performance measurement

This section briefly summarizes the salient features, strengths and limitations of several alternative empirical cost and productivity estimation methodologies from an academic perspective. The reader is referred to Coelli, et al. (2005) and Berg (2010) for more detailed but approachable summaries of performance measurement techniques and their application to infrastructure industries. We focus on three alternative approaches to estimating firm performance, which are the econometric, DEA and non-parametric TFP index number approaches.

A.7.1. Econometric approaches

Econometric approaches rely on the specification of a stochastic parametric production or cost function, thereby providing a fitted functional form capturing the underlying economic determinants of production, such as the relative magnitude of input usage on outputs, or the impact of increased outputs on costs. A general advantage of all econometric approaches relative to DEA and non-parametric index number approaches is the potential to conduct hypothesis tests between alternative model specifications with econometrics, thereby providing greater statistical confidence in the reliability of the model. A general disadvantage is the need to specify a functional form for the cost function. However, while it is sometimes emphasized that this may impose an inappropriate functional form on the structure, the application of fully flexible cost function specifications such as translog specification substantially ameliorates this issue.

Within the econometric approach there is a significant distinction between *average response* models which account for statistical noise, but otherwise model the average relationship between outputs prices and costs, and *frontier* econometric models which are meant to provide estimates of frontier or best practice technology.

Average response econometric approaches

Average response models can be estimated with ordinary least squares regression (OLS), or a variety of well established and broadly accepted conventional econometric techniques employed by economists to provide improvements on OLS estimates. Such methods, when combined with well specified models consistent with economic theory and controls for firm heterogeneity (see above), provide the potential for extremely robust estimates of the expected relationship between a firm's costs and drivers of those costs, with any remaining deviations attributable to statistical noise. However, given the already low number of observations available to Ofwat for cost assessment, the potential for further mergers and the increased degrees of freedom required for the controls for input prices, capital stocks, and operating characteristics suggested above, the potential for further refinements of cross-sectional modelling at company level is limited.[73]

Given the similarly well established theory of estimating average response models with panel data,[74] there is however potential that these techniques can be employed by Ofwat with appropriate panel data sets. Moreover, the simplest application of panel data sets would involve the pooling of data to increase the available degrees of freedom and improve statistical robustness, as suggested by Kumbhakar & Horncastle (2010), who also summarize a series of studies that have employed panel data analysis in the English & Welsh water sector, while noting that none of these studies have been directly employed in benchmarking.

---

[73] Chapter 8 of Coelli, et al. (2005) details the advantages of average response estimation of production and cost technologies and their appropriateness for cost assessment.

[74] See for example, Baltagi (2005).

Kumbhakar & Horncastle's modelling also suggests that such data pooling, with the inclusion of time dummies allowing for technical change provides more reliable estimates than Ofwat's cross-sectional estimates of these four models. Furthermore, they suggest that application of several standard tools of average response econometric estimation with panel data improve the reliability of assessment of Ofwat's four water opex models. Finally, they also strongly suggest that because of cost interactions between these models, the simultaneous estimation of these models is also superior to Ofwat's separate single equation modelling. They therefore reach the conclusion that the use of panel data would allow Ofwat to break its reliance on the cross-sectional corrected ordinary least squares (COLS) models it has employed since 1994, and that it would and should allow assessment of costs at more aggregated levels, thereby eliminating the need for concern with regard to cost interactions between assessed activities. However, while we are highly supportive of Kumbhakar & Horncastle's conclusions, their models only allow for neutral technical change and do not include the controls for intertemporal and inter firm differences in input prices, capital stocks, quality, and heterogeneous operating characteristics, that we have argued for above.

Oxera's submission to the CC as part of the Bristol Water Review, also argues in favour of the use of panel data for regulatory benchmarking. Moreover, rather than replicating Ofwat's four water opex models it adopts the recommendation of Kumbhakar & Horncastle (2010) and assesses all water related opex at the aggregate level rather than with separate models so as to avoid issues related to cost function nonseparability, thereby specifically supporting the recommendations of Section 3.3 above. To this extent we are highly supportive of Oxera's conclusions.

However, the Oxera submission also argues that simple data pooling is appropriate as the hypothesis of parameter change was rejected in the six year panel covering the 2003/4 to 2008/9 period. We believe that both previous work for Ofwat and previous academic evidence calls this result into question. Thus, it is generally accepted that considerable improvements in performance have occurred over this period. More specifically, Ofwat's own assessment of continuing efficiency improvement factors for water opex in the 1999, 2004, and 2009 price reviews suggests that significant changes in the underlying technology, and hence cost model parameters should have occurred over this period. Oxera's finding in favour of a pooled model suggests, however, the absence of continuing efficiency change, or technical change as it is called in normal economic parlance.

We therefore turn to two previous examples of total opex cost modelling. The modelling approach of Bottasso & Conti (2009b) estimates total water opex costs in a quasi-fixed capital stock model, as discussed above in Section A.2. They also employ a translog cost function allowing for fully variable returns to scale, cross firm and intertemporal variation in input prices, and full flexibility in the cost interactions between volumetric outputs, connections, and area output variables.[75] The model also includes appropriate controls for input prices. As such, this modelling approach is by far more consistent with the economic approach to cost modelling detailed above in Sections 3.1-3.4 than the Oxera approach. Moreover, their model allows for

[75] We note that while Bottasso & Conti do not discuss their result, their reported parameters suggest the presence of significant cost interactions between the various water outputs specified in their model, thereby giving further evidence with the regard to cost interactions discussed in Section 3.3.

the potential for parameter change which is consistent with technical change, or equivalently "continuing efficiency change" in Ofwat's regulatory terminology, as discussed in Section 3.5. Furthermore, we note that their approach to parameter change does not assume a simple dummy shifter which would imply neutral or equiproportional reductions in costs associated with technical change. Instead, as suggested in Section 3.5, their approach allows for the possibility that technical change can be output and/or input biased. Thus, for example, the model allows for the reasonable possibility that as technology changes over time, production tends to become more capital and less opex intensive.

In contrast to Oxera, Bottasso & Conti (2009b) find statistically significant evidence of parameter change, and find that technical change was substantial and statistically significant for most years in their 1995/6 to 2004/5 sample. Moreover, in work using very similar specifications to Bottasso & Conti, and commissioned and published by Ofwat, Stone & Webster Consultants (2004b) reach similar conclusions, and find evidence of statistically significant cost function parameter change in both the 1995-2000 and 2000-2003 period for a model of total WaSC opex, and during the 2000-2003 period for their model of WoC total opex. Moreover, the Stone & Webster work demonstrates how considerable changes over time in WaSC and WoC operating characteristics influence assessed costs and hence measured productivity growth.

Thus, we conclude that as an econometric approach employs all the data to estimate the underlying relationship between outputs and costs, data pooling is likely to result in an upwardly biased estimate of costs in the final year of the sample, if technical change is present but not allowed for. The results of Bottasso & Conti, and Stone & Webster demonstrate the need for controls for parameter change, which is consistent with the presence of technical change, and hence the presence of continuing efficiency improvement. It is therefore our considered opinion that the simple pooling of well specified economic models of costs is unlikely to be appropriate for all but exceedingly short panels of two to three years where it might be plausible that evidence of statistically significant technical change will not be found.

However, they also strongly suggest that average response estimation can be employed to provide estimates of the impact of technical change, scale change, productivity growth and other determinants of firm performance on a firm's costs over time, concepts which were all discussed in Section 3.5. It is also our opinion that econometric specifications have an advantage over DEA approaches in this respect, as technical change can be readily allowed for with the addition of a few parameters allowing for trend change in other parameters. This implies that provided sample sizes are sufficiently large, the additional parameters required to allow for technical change will be far less than then number of degrees of freedom provided by additional years of data.

We finally note that the principal potential regulatory weakness of average response approaches is that they do not allow for the estimation of frontier costs, and hence do not provide a direct estimation of inefficiency that is theoretically consistent with the efficiency catch up component of a regulatory X factor. However, we nonetheless wish to emphasize the strong econometric robustness, and potential future application of average response econometric approaches to regulatory cost assessment in the English & Welsh water and sewerage industry. Thus, we believe that development of average response models along the lines of those provided by Bottaso & Conti (2009) and Stone & Webster (2004b) have strong potential to provide superior estimates of firm costs, which might actually reduce the need for regulatory discretion in regulatory cost assessment. However, this would require the availability of sufficient data to allow economically consistent panel estimation.

Frontier econometric approaches

Despite this finding in favour of the potential application of economically robust average response models, Ofwat's primary use of cost assessment is likely to remain the determination of the relative efficiency of firms, so as to both identify inefficiency and properly incentivize its elimination. We therefore turn to consideration of *frontier* econometric approaches, which are fundamentally designed to provide an estimate of cost inefficiency. Nevertheless, there are considerable and important differences between *deterministic* approaches such as COLS, which do not allow for statistical noise, and Stochastic Frontier Analysis (SFA), which attempts to allow for both inefficiency and random error or noise.[76]

COLS has generally perceived benefits related to simplicity and transparency. However, the COLS approach relies entirely on an underlying estimated average response OLS model. Thus the OLS model is "corrected" to form a frontier model, by making the assumption that the estimated OLS residual, which is by definition assumed to consist entirely of statistical noise, can be interpreted entirely as a measure of inefficiency. As Weyman-Jones et al. (2006) argue, simple application of confidence intervals on the predicted values in the resulting COLS model suggests that in many cases we should have little confidence in the resulting estimates of inefficiency. Thus, consideration of the theoretical underpinnings of the COLS approach suggests that estimated inefficiency will exceed actual underlying inefficiency. This likely bias therefore leads to the principle regulatory weakness of COLS modelling, which is the need for substantial regulatory discretion in determining what proportion of estimated COLS inefficiency should actually be attributed to inefficiency. This downside is therefore the reason that Ofwat's current regulatory approach uses a considerable degree of discretion and only applies a portion of estimated COLS inefficiency in setting cost targets. We also note that Ofgem's past practice of using the first quartile firm as benchmark costs is an alternative form of discretionary application that also implicitly acknowledges that the underlying econometric modelling approach is a stochastic OLS regression.

The SFA approach addresses this concern by attempting explicitly to allow for both noise and inefficiency, and simultaneously estimating the cost function parameters as well as several additional parameters needed to estimate the relative importance of inefficiency and noise in the particular model. Thus, we are strongly in agreement with Oxera's recent CC submission suggesting that SFA should, in principle, provide more reliable estimates of the frontier, as opposed to the average response technology, while also allowing for noise, and thereby providing superior estimates of inefficiency in comparison to the COLS approach.

Moreover, the SFA approach should, in principle, also have the advantages detailed above for an average response panel model, while further allowing for the estimation of efficiency change over time. Thus, panel SFA based models can be applied in the regulatory context to decompose cost productivity growth into factors such as technical change, efficiency change, and

[76] See Greene (2008) for a detailed discussion focussed on the cross-sectional estimation of deterministic and frontier efficiency econometric models.

scale change, as suggested by Coelli et al. (2003), and for example demonstrated by Saal, Parker, & Weyman-Jones (2007), with, WaSC data. However, we emphasise that as there are only 10 WaSCS this approach was only feasible given the application of an extended panel database covering 15 years.

However, the advantages of SFA come at the cost of a substantially more complex econometric specification which requires relatively strong assumptions to be made with regard to the statistical distribution of inefficiency.[77] Nevertheless, as recent application of these techniques in German electricity distribution benchmarking and modelling for the ORR demonstrates, SFA can be employed to provide improved estimates of both firm costs and inefficiency in the regulatory context. However, given the need for distributional assumptions and the resulting potential robustness issues, any Ofwat development of SFA modelling – particularly panel SFA – needs to allow for testing against the parameter estimates from less ambitious techniques. In addition, the length, number of available variables, and data consistency of Ofwat's post accounting separation databases will also be fundamental determinants of the feasibility of applying SFA in future cost assessment.

A.7.2. Data envelopment analysis (DEA)

Traditional DEA models rely on linear programming techniques to construct a non-parametric frontier or envelope around the data, thereby providing a fitted production frontier detailing the relationship between a firm's outputs and efficient input quantities. Firm's using inputs in excess of those indicated on the DEA frontier are therefore productively or technically inefficient, with the distance from the frontier indicating the level of inefficiency

Appropriate estimation of cost or allocative efficiency in the DEA framework requires data on both input prices and quantities, and specification of additional constraints to the DEA programme which have the effect of identifying the unique cost minimizing point on the estimated production frontier.[78] Thus, while it is common for many researchers to specify DEA "cost models" by simply employing opex or other measures of cost as an input, this approach is consistent with neither economic nor DEA theory. Thus, as with econometric approaches, it is likely that regulatory DEA based cost assessment will be substantially improved with appropriate controls for input prices, as managers will choose to use alternative combinations of inputs if the relative prices of inputs vary substantially across companies or time.

The principle advantages of DEA relate to its relative simplicity and the intuitive and fully flexible relationship between actual data points and the identified firms (peers) that define the production frontier. The principle disadvantage of the traditional DEA approach is its failure to allow for noise, as in COLS modelling. Moreover, while estimation strategies do allow for the inclusion of operating characteristics as a determinant of costs, it is our opinion that traditional DEA models are far less capable of controlling for heterogeneity between firms than econometric approaches. Recently, extensions of traditional DEA have been developed with the aim of allowing for both noise and operating characteristic heterogeneity. However, while these approaches now have wide academic acceptance, they involve complex modelling which is

[77] These are, however, now available in relatively cheap off the shelf econometric packages such as LIMDEP and STATA.

[78] See for example Thanassoulis (2001).

unlikely to be accepted as understandable and transparent in regulatory application within the UK.[79]

DEA models can also be extended to allow for panel data. One approach is to simply pool the data over multiple years and thereby estimate a common frontier. The DEA pooling approach can be reasonably assumed to provide appropriate estimates of best practice costs in the final year of the pooled sample, even if technical change is present. This is because only the firms with the best relationship between outputs and costs will influence the frontier, and such best practice firms are likely to be drawn from the end of the sample if technology is improving over time. Thus, if one leaves aside traditional DEA's fundamental weaknesses in controlling for both noises and firm specific heterogeneity, the regulatory application of DEA pooling is arguably superior to econometric pooling of panel data, if one's goal is to identify best practice costs in the final year of a panel data. However, given the estimation of a common frontier, the DEA pooling approach does not allow the identification of efficient frontier costs for each time period in the data. Nor does this approach make it possible to distinguish movement towards the frontier (efficiency change) from change in the frontier (technical change).[80]

An alternative approach is to estimate a distinct DEA frontier for each year in the data, thereby allowing for the determination of efficient frontier costs in each year, as well as technical and efficiency change between years. However, this approach amounts to the estimation of separate cross-sectional DEA frontiers in each year, and would not help overcome the issue of estimating costs if the number of cross-sectional observations was limited. In consequence, we believe that there is unlikely to be any strong potential for developing this repeated cross-sectional DEA approach, or even any simple cross-sectional DEA assessment with Ofwat's regulatory data at firm level. We do, however, note the possible exception of the potential application of DEA in sub-company modelling if the number of cross-sectional observations is sufficient. However, the above provisions with regard to potential cost interactions between sub-activities also apply with DEA modelling.

We finally consider sequential DEA modelling of panel data. This approach estimates a series of yearly frontiers, which are estimated for each year by sequentially adding each new year's data to a common data pool and re-estimating the frontier. This approach allows the identification of efficient costs in a given year, and hence allows for the separate identification of both technical change and efficiency change. However, if cross-sectional data is sparse, it would still require the pooling of several years' data before the year of assessment to allow estimation of frontier costs in that year. Thus for example, if only 10 cross-sectional observations were available, and the analyst believed that a minimum of 30 observations were required to adequately model firm costs, a frontier could only be specified for the third and any subsequent years.

In summary, issues related to allowing for statistical noise and controlling for differences in operating characteristics, limit the applicability of traditional DEA approaches in the UK. If we downplay these considerable limitations, simple data pooling in the DEA context can potentially allow reasonably robust estimation of efficient costs in the final year of a sample, which we

[79] See Simar & Wilson (2008), for a discussion of these recent developments in DEA theory.

[80] See Thanassoulis, et al. (2008), for further discussion of cost change in DEA modelling and the decomposition of overall productivity change into technical and efficiency change.

believe would need to be at least three years in length. However, we have also seen that application of more sophisticated panel approaches that allow consistent estimation of frontier change over time still require a relatively large number of cross-sectional observations, or the willingness to apply sequential DEA techniques. These more sophisticated panel data DEA approaches have recently been applied for assessment at the aggregate company level for both WaSCs and Dutch water utilities, and illustrate the academic potential for employing these techniques.[81] Moreover, both papers establish the feasibility of employing more sophisticated DEA techniques to establish longer term trends in the productive performance of companies, and how this influences their profitability and consumer prices. We therefore emphasise their potential application if Ofwat were to move from benchmarking to a yardstick regime in the future. However, given the relatively low number of observations available in both contexts, it is as yet unclear whether the underlying cost estimates provided by such models are sufficiently robust estimates of the efficient costs, as required for regulatory benchmarking. We therefore conclude that while it would be worthwhile to further explore the potential application of DEA approaches, on balance econometric approaches are potentially superior for direct regulatory benchmarking in the English & Welsh context.[82]

However, before actually concluding this section we must note the recent contribution of Bogetoft & Otto (2011) which discusses the application of both traditional and more modern DEA approaches (as well as SFA) to benchmarking. We moreover acknowledge that their book details the apparently successful regulatory application of DEA approaches to energy regulation in Germany and Norway, as well as the development of a benchmarking scheme for Danish water utilities by the Danish Water and Waste Water Association. However, the notable difference in each of these applications is the presence of, by UK standards, a very large number of small utilities. We therefore emphasize that it is this difference in the number of observations available, which results in our less favourable assessment of the potential use of DEA modelling approaches. Fundamentally, this data limitation implies that UK regulators would need to invest substantial regulatory effort in developing and understanding sophisticated DEA panel models before DEA type approaches could be applied.

A.7.3. Non-parametric index number approaches

Non-parametric total factor productivity (TFP) index numbers essentially use input and output prices to respectively provide information with regard to the relative costs of inputs and the relative value of outputs, thereby allowing the construction of aggregate indices of outputs and inputs, from which TFP indices can be constructed.[83] The clearest benefit of this approach is that it can be easily employed to provide relative performance estimates with limited data as only data on input and output quantities, total costs, total revenues and prices are required. In fact the approach requires a minimum of only two firms to allow meaningful cross-sectional performance comparisons between firms, and similarly requires only two years of data for a single firm to consistently assess trends in performance. Moreover, cross-sectional indices

[81] De Witte & Saal (2010) and Maziotis, Saal, & Thanassoulis (2010).

[82] Bogetoft & Otto (2011).

[83] Chapter 4 of Coelli, et al. (2005) provides a general introduction to the economic theory of index numbers and productivity measurement as does Hackman (2008).

provide easily understandable measures of relative productivity between firms, which are inversely proportional to the potential efficiency catch up of the laggard firm relative to the best practice firm. When panel data is available, Maziotis (2010) demonstrates that index numbers can be employed to provide estimates of frontier shift and frontier catch up, and this is feasible even with extremely short panels and a limited numbers of firms. Moreover, Maziotis (2010) suggests that this approach can be readily integrated into models of firm profitability that can be employed to assess the link between trends in productivity, profitability, input and output prices, and consumer prices. Thus, such index number approaches can be seen as a potentially useful tool in assessing company performance even if only a limited numbers of firms are available for assessment.

This ready applicability does however come with some shortcomings. Firstly, non-parametric index number approaches require the availability of reasonable proxies for output prices as well as input prices, thereby making their application infeasible in the absence of output prices. This issue makes their application particularly infeasible in the context of accounting separation, as output prices for each component of a water company's value chain are unlikely to be available in the near future. Moreover, while we have argued above that economically consistent estimation of costs with econometric and DEA techniques should include controls for input prices, estimation with these techniques can still potentially proceed in the absence of some or all input prices. This is simply not the case for index number approaches which require input price and quantity data for computation. Secondly, the approach is relatively inflexible with regard to controlling for differences in operating characteristics and output quality between firms, thereby making appropriate controls for cross firm heterogeneity relatively difficult to implement. Thirdly, the approach assumes the presence of constant returns to scale, which is a significant disadvantage given our above discussion with regard to the appropriateness of allowing for fully variable returns to scale when conducting cross company comparison.

Nevertheless, despite these potential shortcomings, Maziotis (2010) suggests that there is sufficient potential to apply index number techniques at company level, and to gauge relative performance trends between firms and across time. We therefore suggest, that particularly because of their applicability with an extremely low number of observations, Ofwat could potentially employ index number approaches as a readily applicable company level consistency checking tool, with which relative company performance measures derived with econometric or DEA methodologies can be cross checked. Moreover, they also have the strong potential of providing readily understandable, and easily estimated measures of individual company productivity growth performance. We also believe that this latter approach, might particularly complement the more outcomes focused approach to regulation that has recently been under consideration by Ofwat.

A.8. The potential application of panel data models in water regulation

Developing panel models that are consistent with the modelling approaches we have detailed above has the potential expand the quality of regulatory cost assessment. However, the development of sophisticated panel models along the lines of those suggested by Coelli, et al. (2003), would also open up the potential to measure and assess trends in firm productivity growth rates and provide better measurement of the sources of such productivity growth rates.

While it has not been the primary focus of the above academic assessment, many of the academic papers we have considered demonstrate the potential applicability of a variety of sophisticated panel based approaches, such as index number, econometric, and DEA based approaches that have provided alternative perspectives on performance trends in the English & Welsh water industry. These papers include, among others, studies that have analysed technical change (Erbeta & Cave, 2006), industry structure (Stone & Webster, 2004a ; Bottasso & Conti, 2009a; Saal, Arocena, & Maziotis, 2010), total factor productivity growth and its relationship to firm profitability, the relative performance of firms (Saal & Parker, 2001, Maziotis, Saal, & Thanassoulis, 2009) and the total cost saving impacts of privatization and regulation (Saal & Paker, 2000) Moreover, opex productivity growth and the impact of regulation has been analyzed by Stone & Webster(2004b) and Bottaso & Conti (2009b), with the former providing realistic opex productivity growth estimates that effectively served as a check on the opex X-factors set in the 2004 price determination. Similarly, Saal, Parker, & Weyman-Jones (2007) provides a panel based approach to assessing firm efficiency, while also decomposing WasC TFP into technical change, efficiency change, and scale change. This paper particularly illustrates well the potential link between panel econometric models, estimated productivity growth rates, and realized X-factor performance. Consideration of the previous English & Welsh literature as a whole therefore suggests that there is strong potential for applying a variety of alternative methodological panel approaches to regulatory performance assessment. However, we would also note that considerable effort would be required to reconcile these approaches with regulatory requirements.

We would further argue that such sophisticated modelling approaches implement a cost assessment approach that is consistent with the intended cost reducing incentives of a price cap regulation system. This is because price cap regulation, as currently implemented by Ofwat, explicitly seeks to incentivise efficiency by setting regulatory catch up as well as continuing efficiency change X factors. Therefore if such regulation is effective, regulated firms should improve their productivity, and this should be attributable to both improved efficiency and technical change, with the latter term being fully comparable to Ofwat's definition of continuing efficiency improvement. Thus, developing sophisticated panel data models that allow for technical, efficiency, and/or productivity change determination could substantially improve Ofwat's ability to assess both the past achievement and potential future potential of firms to achieve X factor efficiency savings.

There is also considerable discussion in the current regulatory context with regard to reducing regulatory burden and data collection, and the potential application of alternative less intrusive regulation, through, for example, greater reliance on competition, and more reliance on key performance indicators. However, it is likely that for the foreseeable future Ofwat will need to both appropriately incentivize cost efficiency in regulated activities, and ensure that consumer prices are cost reflective in the long term. Given this long term need we wish to briefly speculate with regard to the potential of employing sophisticated panel model approaches in this context.

Consideration of the academic panel models detailed immediately above reveals that they are able to provide reasonably well specified cost models, while employing an insubstantial subset of the entire regulatory databases collected in Ofwat's June Returns. While, considerable further research would be required to ensure that such models could be made consistent with Ofwat's regulatory requirements, the relatively low data burden of these models is suggestive. Thus, we are willing to speculate that if policy makers wish to reduce regulatory burden, there is the potential that appropriate and accurate high level cost assessment of a firm's total regulated costs or its total opex costs could be achieved through the careful application of econometric, DEA, and/or panel data techniques. Thus, for example, such panel models could be employed to calculate realised firm specific relative efficiency and rates of technical change, efficiency change, and or productivity growth which could then be employed to determine X factors consistent with those currently employed by Ofwat.

Alternatively, if policy makers were willing to contemplate the movement to a yardstick regulation regime such as that implemented in Norwegian electricity in 2007, and discussed in Bogetoft & Otto (2011), Ofwat could use determinations of average, best, or median observed ex post costs or productivity growth rates to set X-factors. As Bogetoft & Otto indicate, there are potential pitfalls to this yardstick approach as windfall profits would be retained by the industry, dynamic risks may be passed to consumers, and there is a potential risk of regulatory capture by consumers or industry. Nevertheless, if policy makers were willing to take such risks, it might be worthwhile to consider the potential of further developing sophisticated panel models for cost assessment and regulatory X-factor determination. However, given the already low number of observations in the English & Welsh water industry, this approach would require the continued collection of consistent and relatively long panel data sets similar to those employed in the academic studies detailed above.

We finally speculate with regard to the potential for Ofwat to employ panel data techniques in order to jointly assess the relationship between a firm's cost performance, profitability, and ensuring that consumer prices are cost reflective in the long term. Particularly if we see the introduction of more competition and/or alternative regulatory models, Ofwat might consider the further development and adaptation of firm level profit decomposition techniques such as that illustrated in De Witte & Saal (2010). This paper illustrates a panel performance measurement technique that not only provides a decomposition of the sources of firm productivity growth (including efficiency and technical change), but also assesses its impact on firm profitability and hence consumer oriented allocative efficiency. While this paper is implemented with a DEA based cost assessment, there are also academic examples of such profit decomposition techniques which employ index number and econometric estimation, and which could potentially be developed to satisfy such regulatory requirements.

ANNEX B: COMPETITION COMMISSION CASE STUDIES

In this Annex we set out three case studies of Competition Commission ('CC') decisions which have analysed Ofwat's use of comparators. These case studies note the importance of comparators to Ofwat's ability to set the price caps and set out possible mitigation strategies if a comparator is lost.

B.1. The Bristol Water appeal

Following Ofwat's PR09, Bristol Water plc disputed the price determinations made by Ofwat and on 8 February 2010 Ofwat made a reference to the CC.[84] The full determination process was reviewed and each building block was assessed in turn and adjustments were made by the CC if it considered they were justified. However, the CC did not specifically review the appropriateness of Ofwat's relative efficiency econometric modelling and did not make any recommendations regarding this.

Relative efficiency

The CC reviewed Ofwat's operating cost econometric models and based on the data and outputs of the models concluded that these were sufficiently reliable to estimate inefficiency scores for each company. The CC considered the targets set by Ofwat for Bristol Water against its own analysis, the evidence provided by Bristol Water and evidence submitted by Oxera (see Box A.1 below). The CC decided that, on balance, the weight of evidence was consistent with Ofwat's classification of Bristol Water as being within Upper Band B.[85]

Box B.1: Oxera report on Bristol Water's efficiency (Independent submission to the CC)[86]

The report builds on an earlier paper by Oxera that was submitted to Ofwat as part of PR09. This report reviews the then current methodology for assessing relative efficiency used by Ofwat (COLS) to determine whether this can be improved upon through the use of panel data, different econometric methodologies, and/or model specification. While Oxera consider that there are many aspects of Ofwat's modelling that have much merit, the authors note three issues where their suggested improvements could increase the robustness of the efficiency estimates.

1. that modelling with "panel data is statistically valid and that this results in greater precision";
2. as a panel data set is being used it is now possible to include significantly more cost drivers in the models. Thus, modelling at the aggregate opex level for water services is feasible and may offer a number of benefits, including "removing the requirement for both cost separability and consistency in cost allocation, and allowing for the inclusion of multiple cost drivers; and
3. Ofwat's use of COLS does not correctly account for noise in the modelling. Oxera considered that the use of SFA would be superior as it does not require a subjective adjustment to account for such noise.

With regards to the former part of the first issue above (statically valid), Oxera posit that the data meets the 'poolability hypothesis' i.e. that the parameters (including the intercept and error variance) do not change over time. If this is the case then the error variance can be estimated more precisely and this in turn lessens predictive uncertainty. Oxera show this to be the case for its two models (total opex, and total opex less special factors) by rejecting the hypothesis that there is structural

[84] Competition Commission (2010).

[85] Competition Commission (2010), p58.

[86] Oxera (2010).

change. In regards to the later (precision), Oxera show that the confidence intervals are significantly small when panel data is used compared to cross-sectional.

In justifying the use of total opex, Oxera note that modelling at this level avoids the issue of cost separability, and reduces the impact of cost allocation differences between the companies.

Finally, Oxera propose that SFA be used instead of COLS as SFA is able to decompose the residual term into inefficiency and noise (something COLS cannot do and requires a subjective judgement on the level of noise in the data). Oxera do note that one disadvantage of SFA is that it often requires a large data set.

As the appeal was in relation to Bristol Water's price cap the CC made no comment on the appropriateness of Ofwat's modelling (aside from accepting their findings) or the use of panel data.

B.2. The Competition Commission on South East Water and Mid Kent Water

In 2007 the CC carried out an examination of the merger between water-only companies South East Water and Mid Kent Water.[87] The final report focused on the question of how the loss of a comparator affects Ofwat's ability to effectively regulate the sector. The CC considered Ofwat's approach to comparative analysis, the structure of the water/ water and sewerage sector, and any possible solutions that may be applied to mitigate any negative impact from the merger.

The CC noted that the loss of a comparator could affect Ofwat's ability to carry out comparative analysis in one of four ways:

1. it could result in the loss of a benchmark company in Ofwat's econometric models;
2. it may reduce the precision of Ofwat's econometric models;
3. it may affect Ofwat's ability to make cost-base comparisons to challenge cost-base estimates; and
4. it may affect Ofwat's ability to make qualitative comparisons.

The CC considered a range of options that might be available to Ofwat to overcome the loss of a comparator. These included:

1. relying on additional data (such as sub-company and panel data, or international or cross-sector comparators);
2. utilising stochastic frontier analysis (SFA) and/ or data envelopment analysis (DEA); and
3. adopting simpler statistical approaches and qualitative comparisons, similar to those Ofgem rely on.

The CC noted that it would be difficult and costly to collect robust sub-company data and that the integrated nature of a water network would make it difficult to allocate costs to activities accurately. On panel data, while noting its usefulness and the fact that it is already available, the CC warned that its utilisation requires the cost structure in the industry to have remained constant over time or be relatively simple to model.

The CC noted that cross-sector data could be used for a range of standard activities, such as customer service, and that while it was difficult to collect consistent and comprehensive

[87] Competition Commission (2007).

international data, owing to the different regulatory regimes that are in place, this data could still serve as a useful cross-check.

The CC argued that both SFA and DEA require large numbers of comparators in order to be robust, while they were also likely to be just as sensitive to the loss of a comparator as Ofwat's current econometric models.

B.3. The Monopolies and Mergers Commission on Severn Trent and South West Water

The Monopolies and Mergers Commission (MMC) investigated the proposed merger of Severn Trent and South West Water, two water and sewerage companies, in 1996.[88] The MMC's investigation focused on the impact of the loss of an independent comparator on Ofwat and the Director General of Water Services' ability to carry out comparative analysis and generate effective comparative competition in the water and sewerage sector. It also considered the value of the loss of South West Water as a comparator, how it compared to the expected benefits of the merger, and any potential remedies. The MMC concluded that the loss caused by the merger would be too big and that no effective remedy was available, so it recommended that the merger should be prohibited.

The MMC noted that the number of companies required for comparative analysis varies according to the type of analysis performed, from direct comparisons of specific costs between two similar companies, to econometric models that requires a large number of comparators in order to provide a useful degree of precision. The MMC identified four ways in which the loss of a comparator could adversely affect effective comparative analysis:

1. the loss of a potential benchmark if the company is at or near the efficiency frontier;
2. reducing the confidence with which the regulator carries out its activities owing to fewer data points and possibly varying quality of information;
3. efficiency and service quality could be affected through reduced cost competition between companies; and
4. the way the Director General of Water Services carries out other tasks could be affected by the loss of a comparator.

The MMC also noted that the loss of each additional comparator would have a greater adverse effect on the regulator's ability to make comparisons.[89]

[88] Monopolies and Merger Commissions (1996). It should be noted that a separate bid by Wessex Water for South West Water was assessed at the same time. Obviously the same approach to analysis was undertaken.
[89] Ibid, para 4.37-4.38, p. 57.

ANNEX C: REGULATORY CASE STUDIES

In this annex we set out a number of regulatory case studies:

- Ofwat;
- Ofgem;
- Postcomm;
- ORR; and
- the New Zealand Commerce Commission.

These case studies predominantly focus on operating cost comparative analysis. The regulators on which the case studies are based tend to rely on relatively simplistic unit cost or financial/ volume ratios for capex benchmarking – this can be contrasted with the more formal and data intensive cost base approach that Ofwat has used for investment. As discussed in Section 4.2.2, these benchmarking approaches generally require data that is more readily available to companies (or at least better understood), are applicable for sub-company analysis and the methods require few (if any) controls, therefore we do not consider them in great detail here.

Regulator (country)	**Ofwat (UK)**
Sector (i.e. Water, Gas, Electricity)	**Water (PR09 relative efficiency assessment)**
Elements of the regime	**Description of current approach**
Market structure (e.g. degree of competition, business levels, etc)	21 licenses, one each for the 21 water companies (of which 10 are water and sewerage companies and 11 are water-only companies) There is limited competition in the form of new appointments (for a defined area) and water supply licenses, the latter are available to large consumers only. The Water Act 2003 had a provision (enabled from 2005) to allow competition through the supply and carriage of bulk water, however this has seen limited results.
Unit of analysis (e.g. group, licence level, sub-business, activity)	Capital efficiency assessment was undertaken using the cost base comparative tool to compare company estimates of capital works unit costs for a representative range of standardised capital projects (standard costs). Ofwat undertook operating efficiency assessment separately for water and sewerage services using econometric and unit cost models. For each service i.e. water and sewerage, there are sub-models based on the activities undertaken by water companies. Water service was divided into 4 econometric models: 1. water distribution; 2. water resources and treatment; 3. water power; and 4. water business activities. Sewerage service was divided into 2 econometric models and 3 unit cost models. The econometric models were for: 5. large sewage treatment works; and 6. sewerage network, including power.

	The unit cost models were for: • small sewage treatment works; • sludge treatment and disposal; and • sewerage business activities. Data was collected from June Return 2009. Different variables were used for each of the econometric models. They are(numbered items correspond to list of econometric models above): 1. Length of mains/number of connected properties 2. Number of sources/distribution input and proportion of supplies derived from boreholes 3. Average pumping head and distribution input 4. Number of billed properties 5. Total load, activated sludge, BOD_5 6. Sewer length; area of sewer district; resident population; and holiday population.
Time frame (e.g. current year, historic [panel, pooled], forward looking model)	For opex Ofwat used cross-sectional data from the June Return 2009 (i.e. 2008-2009 data). Ofwat used sub-company data for the opex models.
Regulatory mechanism (e.g. deterministic, discretionary)	For opex the results of the modelling at each functional level were combined and adjustments were made for atypical opex costs, company-specific factors, leakage allocation (water only), cross-subsidies, and pension costs. Each company was assessed relative to the benchmark company and ranked into bands A to E, where A is the most efficient. The banding was used to assign an efficiency factor while setting price limits.
Analytical technique used (e.g. unit cost benchmarking, DEA, SFA, OLS, COLS, etc)	For opex Ofwat used Corrected Ordinary Least Squares (COLS) with some additional adjustments for its assessment of opex.
Any other relevant factors	

Regulator (country)	Ofgem (UK)
Sector (i.e. Water, Gas, Electricity)	**Electricity (DPCR5 cost assessment)**
Elements of the regime	**Description of current approach**
Market structure (e.g. degree of competition, business levels, etc)	14 licences (one each for the 14 Distribution Network Operators – DNOs) owned by 7 groups; of which one group owns three licensees, five groups own two licensees and one group has a single licensee. No competition exists between the distribution networks.
Unit of analysis (e.g. group, licence level, sub-business, activity)	For capex, Ofgem relied on simple unit cost analysis and ratio measures to compare DNO performance. For example, non-load related expenditure was compared using unit costs for different asset types and volume based on age-profiles. Ofgem used three levels of modelling for operating costs - Top-down (operational costs), Single Groups and Groups **Top down** regressions for all Operational Costs **Single Group:** Indirect activity costs were included in a single regression and Network Operating Costs were split into four regressions: 1. Low voltage (LV) & High voltage (HV) Underground Faults (including services), 2. LV & HV Overhead Faults (including services), 3. Inspections and Maintenance, and 4. Tree Cutting **Groups:** Network Operating Costs were disaggregated as per Single Groups and the Indirect Activities are further disaggregated into: 1. Group 1: Network Design, Project Management, System Mapping 2. Group 2: Engineering Management, Control Centre, Call Centre, Stores, H&S and Operational Training, and 3. Group 3: HR and Non-Operational Training, Network Policy, CEO, Finance and Regulation, IT and Property Management. Ofgem undertook the Group 3 analysis at the DNO group level (i.e. parent company level, for example, all EDF's costs summed together) rather than at individual DNO level resulting in 28 data points compared to 56 data points used for other groups Ofgem collected data via the Forecast Business Plan Questionnaires and some data from the Regulatory reporting packs. The number of variables (drivers in this case) used in the regression analysis was seven. They were: 1. Number of Faults 2. Length of Cable Replaced 3. Asset Work Hours 4. Spans Cut & Spans Affected 5. Network Investment (Labour and Contractor costs only) 6. MEAV 7. Total Direct Costs
Time frame (e.g. current year, historic [panel, pooled], forward looking	Ofgem used panel data for a 4 year time frame covering 2005-06 to 2008-09. Regressions were run on the 4 years of data to determine model output costs in 2008-09. The model output costs were compared to the DNOs'

model)	own costs in 2008-09 to determine the overall efficiency scores. The scores were applied to the DNOs own costs to determine 'efficient' costs in 2008-09. The 2008-09 efficient costs were rolled forward into the DPCR5 period to determine Operational Cost baselines.
Regulatory mechanism (e.g. deterministic, discretionary)	The results of the modelling were weighted together and efficiency scores were calculated and given to each of the 14 DNOs. Separate scores were given to Network Operating costs and Indirect costs.
Analytical technique used (e.g. unit cost benchmarking, DEA, SFA, OLS, COLS, etc)	Ofgem used an OLS estimation method for opex; with 19 sets of analyses undertaken; comprised of three core analysis and 16 alternatives (where a single cost item was either added or excluded from the core ones) Network Operating Costs and Indirect Costs were benchmarked to the upper third and upper quartile level of efficiency.
Any other relevant factors	

Regulator (country)	Postcomm (UK); report by LECG
Sector (i.e. Water, Gas, Electricity)	**Mail (Royal Mail Price Control 2005)**
Element ("axis")	**Description of current approach**
Market structure (e.g. degree of competition, business levels, etc)	Royal Mail is the only company regulated by Postcomm, however Royal Mail has separate mail centres and delivery offices that can be considered as discrete units for the purpose of internal benchmarking purposes. Royal Mail maintains market power in most parts of the market; however there has been entry into some areas of the market which has increased competition. Postcomm is currently investigating the scope for increasing competition in the sector. Following its November 2010 consultation, Postcomm noted that in markets where it had identified growing competition, they intend to give Royal Mail greater commercial freedoms and are therefore proposing the following deregulation: • Substantial deregulation of packets and parcels weighing more than 2 kilograms • Retail price controls will be removed from all packets and parcels weighing more than 500 grams, and in parts of the pre-sorted bulk mail market • Royal Mail will be given greater freedom to compete in the pre-sorted bulk mail market, where the company has lost significant volumes to competitors Since Royal Mail offers various services, it operates under a number of price control regimes, some of which include maximum caps on allowed revenue in some areas, while there are headroom controls requirements that restrict its ability to increase prices of some of the access products.
Unit of analysis (e.g. group, licence level, sub-business, activity)	LECG was engaged by Postcomm to determine the scope for efficiency savings in Royal Mail. They did this through both top-down and internal efficiency benchmarking. The internal efficiency benchmarking focused on sub-company operating cost data, namely the mail centres and delivery offices. A number of different approaches were used for the top-down analysis, these included; reviewing and comparing separate historical periods of Royal Mail's performance; a comparison to efficiency targets set and achieved in other regulatory sectors; comparisons of TFP ratios to other sectors; and looking at trends across other international postal operators. For the internal benchmarking of operating costs, data from all royal mail offices was used: 70 mail centres and 1383 delivery offices – some offices were however dropped for reasons of poor data quality. Numerous explanatory variables used as part of the analysis, including: • Number of delivery points • Percentage of delivery points that are businesses • Weighted volume of mail per delivery point • Length of road per delivery point • Delivery zones: major city centres, urban, suburban, etc.. • Mail reduction – i.e. redirections • Mail walk sorted

	• Average distance between delivery office and mail centre • Number of sorting frames • Number of available vehicles at delivery office • Variations in input prices covered by variation in wage rate paid by delivery office • Competitiveness of local labour market/labour force average quality • A quality of service measure, capturing the percentage of all due mail delivered on time.
Time frame (e.g. current year, historic [panel, pooled], forward looking model)	Only cross-sectional data for 2003/4 was used for the analysis
Regulatory mechanism (e.g. deterministic, discretionary)	Postcomm adopted a discretionary approach to using the findings from the efficiency analysis; particularly important given the range of approaches used. Postcomm applied 20% or 10% discount factors to saving implied by deterministic and stochastic frontier analysis DFA and DEA and presented resulting figures. The top 10% of offices, identified through internal benchmarking, were used as the benchmark.
Analytical technique used (e.g. unit cost benchmarking, DEA, SFA, OLS, COLS, etc)	LECG used a range of different techniques. These included: 1. For internal benchmarking: data enveloping analysis (DEA); and both DFA and SFA approaches. 2. For top-down benchmarking: TFP ratios were used.
Any other relevant factors	

Regulator (country)	ORR (UK)
Sectors	Rail (PR08)
Element ("axis")	**Description of current approach**
Market structure (e.g. degree of competition, business levels, etc)	Network Rail is a single national monopoly and there are no direct domestic comparators. It is responsible for the UK's rail infrastructure including tracks, most stations and the signalling network. The Office of Rail Regulation (ORR), measures its cost efficiency in comparison with other international operators in the European Union.
Unit of Analysis	The ORR assesses efficiency across all of Network Rail's activity and expenditure: support functions, operations, maintenance, renewals, and enhancements. They then calculate an overall cost efficiency level. Network Rail overall cost efficiency is compared with the level of efficiency of its European peers. This is obtained by doing top-down analysis on all cost drivers. However, bottom-up analysis is also considered, but used mostly for engineering or operational based analysis. ORR uses the "lasting infrastructure cost benchmarking" (LICB) dataset developed and maintained by the International Union of Railways (UIC). This covers 14 European rail infrastructure managers, including Network Rail, covering the 11 years from 1996 to 2006. The dataset contains maintenance expenditure, renewals expenditure and cost drivers. To assess the scope for improvement in maintenance and renewals (M&R) efficiency ORR conduct econometric efficiency analysis using the UIC's LICB dataset. For operating expenditure, they looked at the long run trends in real unit operating expenditure (RUOE) across a range of UK regulated companies. They initially modelled M&R costs separately, and then modelled the two together. They conclude that their preferred model is the one based on total M&R cost because it means that both the trade-offs between M&R, and any accounting differences between countries in the way in which they record maintenance and renewal costs, were taken into account. ORR's preferred model considers total maintenance and renewal expenditure (cash cost) to be explained by route km (network size), passenger train density (measured as passenger train km on the main line network), freight train density (measured as freight tonnage on the main line network), the proportion of single track on total track km, and a time variable to capture technological progress. • Cost items were mainly maintenance and renewal. • 6 output data: o Freight train km; Freight tonne km; Total tonne km; Total train km; Passenger tonne km ; Network features data: Proportion of track electrified; Number of switches per; track km; Stations per route km; Main track or route km; Ratio of single track to track km
Time frame (e.g. current year, historic [panel/pool], forward looking model	Data used is obtained from the LICB dataset. Only 12 out of the 14 European rail managers data mentioned above were used for the analysis. Two were dropped because of data incompatibility reasons. The data covers 11 years.
Regulatory mechanism (e.g. deterministic, discretionary)	ORR makes discretionary adjustment for Railtrack/Network Rail's renewal cost figures, before and after the Hatfield derailment when computing Network Rail's "steady state" level.

Analytical technique used (e.g. unit cost benchmarking, DEA, SFA, OLS, COLS, etc.)	The final preferred econometric model is SFA with a Cobb Douglas production function. However, OLS was used at an earlier stage in order to determine explanatory variables which were the most significant. The final model was cross-checked against COLS and an attempt was made with DEA.
Any other relevant factors	• Three aspects of efficiency considered: catch up efficiency, frontier efficiency and input prices. • Following PR08, it was agreed with Network Rail that they would work with ORR to conduct further work to explain the drivers behind the maintenance and renewals cost efficiency gap. The analysis was to potentially include a wide range of factors such as technologies and working methods, network/infrastructure configuration, wage rate differentials, differences in geography, macroeconomic factors and differences in government policy - "The ability of Network Rail to control for these different factors will vary, as may the timeframe over which change can be made." • "[S]ub-national level data from five railway infrastructure managers in Europe and North America (including Network Rail) collected directly from the infrastructure managers, was used for separate analysis in PR08, a previous report]"

Regulator (country)	New Zealand Commerce Commission (New Zealand)
Sector (i.e. Water, Gas, Electricity)	Electricity distribution (Price control 2004)
Element ("axis")	**Description of current approach**
Market structure (e.g. degree of competition, business levels, etc)	The New Zealand electricity market is split into three vertical segments: Generators (or suppliers); distribution; and retail. While, a company is allowed to own a generation business and a retail business, is not allowed to own distribution operators i.e. distribution need to be separate. Distribution operators are made up of private sectors owners and municipalities. Companies can own more than one operators, but most operators a independent.
Unit of analysis (e.g. group, licence level, sub-business, activity)	NZ Commerce Commission engaged Meyrick (an economic consultancy) to conduct an assessment of the scope for cost efficiency gains. Meyrick's approach was to decompose the X factor into two components: a 'B' factor reflecting the overall or average productivity trend for electricity lines businesses and two 'C' factors reflecting relative productivity and profitability of each distribution business. These factors then formed the base for calculating an efficiency index. Meyrick used data covering 29 distributors over the period of 1996 to 2002. These distributors are grouped into two categories: urban and rural. They are also grouped into three classes: earning high, average and low rates of return and allocated C factors of -1, 0 and 1 percent respectively. Variables used in the analysis: • Outputs variables: throughput, system line capacity and connection numbers. • Inputs are broken into five categories: operating expenses, overhead lines, underground cables, transformers and other capital. • Variables mentioned in report: energy delivered in kilowatt hours, system line capacity in MVA kilometres and connection numbers, underground lines capital, transformer capital and other capital items.
Time frame (e.g. current year, historic [panel, pooled], forward looking model)	The data used by Meyrick was historical, covering the period from 1996 to 2002
Regulatory mechanism (e.g. deterministic, discretionary)	Meyrick calculated various comparable multilateral Multilateral Total Factor Productivity (MTFP) indices for each of the distributors; distributors are also classified into three groups with "C" indexes. • "Those distributors performing better than the industry average on productivity levels and those earning low rates of return would be set less onerous overall X factors compared to those performing near the industry average. Those performing worse than the industry average on productivity levels and those earning high rates of return would be set more onerous overall X factors compared to those performing near the industry average. These comparisons should ideally take account of differences in distributors' operating environments to the maximum extent possible." A relatively deterministic approach was used with each operators being placed in a band (1, 0, -1) depending on the results of the modelling for each C-factor.

Analytical technique used (e.g. unit cost benchmarking, DEA, SFA, OLS, COLS, etc)	The consultants used MTFP and estimated: • "B" and "C" comparable indexes are produced for each distributor. • MTFP – only using C factor , and both B and C factors are looked at in an integrated framework.
Any other relevant factors	Two efficiency targets could potentially be considered by the regulator: (i) the absolute level of the distribution operator relative to best practice (i.e. a frontier); and (ii) a distribution operator's growth rate relative to competitors.

ANNEX D: ACADEMIC CASE STUDIES

This section provides a few specific academic case studies where panel data has been applied. In Table D.1 we provide a summary of the specific applications of frontier efficiency and productivity measurement, these include numerous panel data models. The D.1 summary table is copied directly from a 2010 discussion paper, *A review of frontier approaches to efficiency and productivity measurement in urban water utilities*, by Andrew Worthing.

Academic (paper)	Determining the contribution of technical change, efficiency change and scale change to productivity growth in the privatized English and Welsh water and sewerage industry: 1985–2000 *By David S. Saal, David Parker and Tom Weyman-Jones, 2007*
Sector	Water and Sewerage
Description of approach	

The paper estimated a quality-adjusted input distance function with stochastic frontier techniques in order to estimated productivity growth rates for the period 1985–2000 – productivity was decomposed so as to account for the impact of technical change, efficiency change, and scale change.

The paper does not provide comparable efficiency levels for each company, but looked at the industry in general.

Both output and input variables were incorporated into a function that enabled the authors to obtain an efficiency distance relative to the frontier.

- Inputs variables used were: full time equivalent employee; operating costs less current cost depreciation, infrastructure renewal expenditures and non-capitalized manpower costs; estimated capital stock
- Output variables used were: connections with water customers; connections with sewerage customers; physical water supply; physical sewage treatment load.
- Exogenous operating characteristics variables used were: water abstractions from underground sources; the ratio of trade effluent loads to resident population; bathing water intensity; the proportion of connected water properties that were metered.

The paper used data obtained from WaSCs covering the period of 1985 to 2000. The authors estimated total factor productivity and decomposed this into efficiency change, technical change return to scale effect.

The model used a translog functional form and a Malmquist productivity index was computed. The error term was decomposed into components: one representing randomness (and is assumed to be independently and identically distributed) and the other measuring inefficiency (and is assumed to be drawn from an independent half-normal distribution that is truncated at zero).

The paper does not compare operators; only changes to the average industry performance are provided since this paper attempts to measure the impact of privatisation on the industry in general.

The paper is an empirical study aiming to evaluate the effects of privatisation and the introduction of a new regulatory regime on efficiency in the water and sewerage sector. Although this is the main focus of the paper, efficiency scores of different operators were estimated while doing so.

Academic (paper)	Improving the econometric precision of regulatory models *By Subal C. Kumbhakar and Alan P. Horncastle, 2010.*
Sector	Water

Description of approach

21 licenses, one each for the 21 water companies (of which 10 are water and sewerage companies and 11 are water-only companies)

The paper examines the impact of the proposed merger of two operators in relation to Ofwat ability to produce accurate estimates for the purpose of benchmarking.

The paper reviewed the then current methodology used by Ofwat (COLS) to determine whether a particular merger would impede on its ability to accurately estimates comparable efficiency values and therefore its ability to appropriately regulate the water industry. This is then followed with recommendations on the appropriateness of econometrics techniques for such tasks. For this particular task, panel data models were judged to be the most suitable.

The paper focuses on Ofwat's opex efficiency analysis – but the authors noted that this could be applied to capex, too.

The authors used data obtained from Ofwat covering the period from 1997/8 to 2007/8 for their analysis.

The authors argued that joint estimation of all the sub-models using the 'seemingly unrelated regression' (SUR) procedure in a cross-section and/or panel data framework can dramatically improve the accuracy of the modelling.

The authors presented measurement improvement that can be obtained using panel data analysis on resources and treatment, distribution, power and business activities.

Two methods were presented: the authors use panel data covering multiple years and cross-section only covering 2007/8. Data used for the paper is historical as follow:

- panel data for 1997/8 to 2007/8 and 2000/1 to 2007/8;
- pooled OLS for 1997/8 to 2007/8 and 2000/1 to 2007/8; and
- cross-section for 2007/8

The authors recommended using panel data instead of simple OLS as panel data produces more accurate estimates necessary to inform adequate decision making when faced with proposed merger(s).

Both the panel data model and the cross-section models subjected to the merger and non-merger of two of the companies show that panel data produces more accurate estimates.

Academic (paper)	Improving the econometric precision of regulatory models *By Subal C. Kumbhakar and Alan P. Horncastle, 2010.*
Sector	Water

Description of approach

The paper examines efficiency measurements issues, applied to the English and Welsh water and sewerage industry. It focuses on separating regulatory and environmental impacts from managerial inefficiency.

The authors plot trends of capex allocated distortion, other expenditure allocative distortion and Labour input distortion.

The dataset used covers the period from 1992/93 to 2004/05 for the ten Water and Sewerage Companies:

- Four outputs variables : the total volume of delivered potable plus non potable water; the total number of household and non-household water service-connected properties; the total number of household and non-household sewerage service-connected properties; the physical amount of waste water.
- Three inputs variables: labour, other operating expenditures and capital.
- Other non controllable variables such as environmental variables were also included

The authors use a two-stage DEA model to obtain measures of managerial inefficiency, separately from general noise and environmental impact.

1. Measures of both technical and allocative efficiencies are calculated.
2. Comparisons of the various operators not included; only minimum and average levels of efficiency are provided.

Academic (paper)	Performance Measurement in the Australian Water Supply Industry *By T. Coelli and S. Walding, June 2005*
Sector	Urban Water Businesses

Description of approach

Various government-owned businesses provide water supply services to Australian customers. Water pricing regulation is carried out by independent bodies in the each state or territory. A generalised CPI – X is represented in the paper as the type the regime adopted by regulators.

The authors used panel data on the 18 largest Australian water services businesses, observed over an eight-year period from 1995/6 to 2002/3, to measure the relative efficiency and productivity growth of the Australian water businesses. They modelled this using a top-down approach, incorporating two output variables and two input variables – i.e. opex and capital.

- Output variables: Number of properties connected and Volume of water delivered.
- Input variables: Operating expenditure (OPEX) and Capital (CAP).

The authors used historical data points in their modelling and a DEA methodology. The following indexes were calculated for each operator:

- Malmquist TFP (an index that measures the TFP change between two data points by calculating the ratio of the distances of each data point relative to a common technology).
- Both technical efficiency and scale efficiency indexes – an additional index was obtained from the product of these two.

The authors also calculated: technical efficiency change (TEC); technical change (TC); TFP change (TFPC); and the potential X factors values.

Regulator (country)/ academic (paper)	The Impact of Regulation on Cost Efficiency: An Empirical Analysis of Wisconsin Water Utilities. *By C Aubert and A Reynaud, 2005*
Sector (i.e. Water, Gas, Electricity)	**Water Utilities in Wisconsin**

Description of approach

The authors noted that firms were not all regulated in the same way in Wisconsin, stating that during the same year, some could be under a regime close to price cap, whereas others could be under another one closer to rate-of-return regulation:

1. The regulated company could ask for a price in any given year. In this case the rate of return regime applies when determining whether to approve the request or no.
2. The regulated company might have decided not to ask for a price increase. In this case the price cap in implementation in the previous year applies.

The authors used a panel of 211 water utilities observed from 1998 to 2000 to show that their efficiency scores could be partly explained by the regulatory framework (price cap or rate of return).

Efficiency levels, although not presented, are obtained by regressing the log of variable cost of each company on output produced and all variable input costs, using stochastic frontier analysis (SFA).

- The variable cost of the utility was the sum of expenses for labour, energy, chemicals, operation supplies and expenses and maintenance
- Two outputs variable were used – the volume in thousands of gallons (Mgal) sold by the water utility to final customers and the number of customers served by the service
- Three technical variables were used – dummies for water utilities that purchase water from another utility, those that use surface water, and the average depth of pumping wells.

The mechanism in place could be considered deterministic in the sense that the regulated company was aware of the conditions set for each type of regulatory regime as mentioned above. The authors used Stochastic Frontier Analysis with a translog function.

This is an empirical paper aiming to assess the impact of the regulatory environment on cost efficiency. However, in to order achieve this, calculating the of level efficiency of each firm are made but not presented – only means, medians of groups are included. Comparison between operators is therefore not included.

Table D.1: Selected applications in urban water utilities

Author(s)	Methodology[a]	Sample[b]	Specification[c]	Technique	Findings
Norman and Stoker (1991)	DEA	28 water-only companies, England and Wales, 1987/88.	Inputs: Manpower, power, chemical and others costs (including an allowance for capital renewal) Outputs: Potable water, properties supplied, average pumping head, length of mains, average peak. Input-orientated CRS.	Descriptive analysis.	Output quantities largely fixed, need to define measures of service quality.
Lambert, Dichev and Raffiee (1993)	DEA	238 public and 33 private utilities, US, 1989.	Inputs: Labour, energy used, materials used, Outputs: Capital value Wholesale and retail water delivered; Input-orientated VRS.	Descriptive analysis.	No significant differences in scale efficiencies between private and public utilities. Most inefficiency results from the overuse of capital.
Bhattacharyya, Harris, Narayanan and Raffiee (1995)	SFA	190 public and 31 private utilities, US, 1992.	Dependent: Variable costs. Independent: Volume of water; energy, labour, materials; water input produced or available for delivery, stock of capital; water input source (surface, ground, both), system loss, age of distribution pipelines, number of emergency breakdowns, length of distribution pipeline, customer type (residential or commercial).	Descriptive analysis.	Cost inefficiency higher in private utilities. Cost inefficiency also positively correlated with size and that major influence on cost inefficiency is breakdowns.
Cubbin and Tzanidakis (1998)	DEA	29 companies, England and Wales, 1992/93	Inputs: Operating expenditure. Outputs: Water delivered, length of mains, proportion of water delivered to non-households. Input-orientated CRS.	Descriptive analysis.	Regression analysis and DEA both suitable for measuring efficiency in water utilities.
Thanassoulis (2000)	DEA	21 water and sewerage companies, 10 water-only companies, England and Wales, 1992/93	Input: Operating expenditure. Outputs: Number of supply connections, length of main, amount of water delivered, measured water, unmeasured water, expenditure on volume. Input-orientated VRS.	Descriptive analysis	Comparison of DE measures of efficiency with efficiency estimates provided by industry regulator.
Anwandter and Ozuna (2002)	DEA	110 water utilities, Mexico, 1995.	Inputs: Personnel, electricity, materials, chemicals, outside services, other costs, specific wastewater treatment costs. Outputs: Water supply, primary treatment, secondary treatment. Nondiscretionary inputs: Water losses (proxy for age of capital stock), population density, nonresidential users. Input-orientated VRS.	Descriptive analysis, second-stage regression.	Decentralisation to the municipal level and appointment of autonomous regulator had no positive influence on efficiency in the absence of competition reform.

Author(s)	Methodology[a]	Sample[b]	Specification[c]	Technique	Findings
Estache and Rossi (2002)	SFA	50 water companies in 29 Asia-Pacific countries, 1995.	Dependent: Operational costs. Independent: Average salary, number of clients, daily production, number of connections, population density in area served, percentage of water from surface sources, number of hours of water availability per day, percentage of metered connections, qualitative treatment variables (chlorination, desalination)	Descriptive analysis.	Cost efficiency not significantly different in private and public sector utilities.
Thanassoulis (2002)	DEA	10 water and sewerage companies, England and Wales, 1994	Inputs: Operating expenditure. Outputs: resident population, length of sewer pipes, size of area served, capacity of pumping in sewerage network. Input-orientated CRS.	Descriptive analysis.	Highlighting of generic influences on efficiency measurement and use of comparative measures.
Bottasso and Conti (2003)	SFA	10 water and sewerage companies, 12 water-only companies, England and Wales, 1995–2001	Dependent: Operational expenditure. Independent: Water delivered, price of labour and capital. Explanatory: Sewerage dummy, length of mains, average pumping head, proportion of river sources on total water sources, population density, volume of water introduced into the distribution system.	Descriptive analysis.	Operating costs inefficiency has decreased over time with inefficiency differential between firms narrowing. Technical and structural requirements impact on cost efficiency.
Tupper and Resende (2004)	DEA	20 Brazilian water and sewerage utilities, 1996–2000	Inputs: Labour costs, operational costs, capital costs. Outputs: Water produced treated sewerage, population served-water, population served-treated sewage. Output-orientated VRS.	Descriptive analysis, second-stage regression.	Network densities and accounted-for water ratio influence efficiency.
Woodbury and Dollery (2004)	DEA and MI	73 water supply authorities, New South Wales, Australia, 1999–2000	Inputs: Management, maintenance and operation, energy and chemical, and capital replacement costs. Outputs: Number of assessments (services to properties), annual water consumption, water quality index (compliance with chemical and physical requirement and microbiological requirements, water service index (water quality complaints, service complaints and average customer outage). Nondiscretionary inputs: Population, properties per kilometre of main, location, rainfall, percentage residential, unfiltered water, groundwater.	Descriptive analysis, second-stage regression.	Technical inefficiencies more substantial than scale inefficiencies. Need for inclusion of service quality outputs.
Aubert and Reynaud (2005)	SFA	211 water utilities, Wisconsin, 1998–2000.	Dependent: Variable costs. Independent: Volume of water sold, number of customers, price of labour and electricity, amount of capital, dummies for water purchased, surface water and average pumping depth.	Descriptive analysis.	Efficiency scores partly explainable by regulatory framework.
Fraquelli and Moiso (2005)	SFA	18 territorial regions, Italy, 1975–2005	Dependent: Total costs. Independent: Network length, number of employees, population served, ratio of population to network length, labour, electricity, materials, services and capital costs.	Descriptive analysis.	Inefficiency partly explained by network characteristics.

Author(s)	Methodology[a]	Sample[b]	Specification[c]	Technique	Findings
Coelli and Walding (2006)	DEA and MI	Australia, 18 water services businesses, 1995/96 to 2002/03	Inputs: Operating and capital expenditure. Outputs: Number of properties connected, volume of water delivered. Input-orientated CRS.	Descriptive analysis.	Need for improvement in specification of capital and provision of water industry price deflators.
Erbetta and Cave (2006)	DEA	10 water and sewerage companies, England and Wales, 1993–2005	Inputs: Number of household and non-household water connections, number of household and nonhousehold sewerage connections, physical amount of wastewater, labour, other operating expenditure, capital expenditure. Outputs: Volume of delivered potable and non-potable water. Nondiscretionary inputs: Water losses, water population density, sewerage population density, time trend, regulatory change dummies. Input-orientated VRS.	Descriptive analysis, second-stage regression.	Regulatory change promoted reduction in technical inefficiency. Price-cap regulation brings inputs closer to their cost-minimising level. Environmental factors influence observed efficiency.
García - Sánchez (2006)	DEA	24 Spanish water utilities, 1999.	Inputs: Staff, treatment plants, delivery network. Outputs: Water delivered, number of connections, chemical analyses performed. Nondiscretionary inputs: Population, persons per household, municipal area, tourist index, average temperature, income, area of greenbelts, economic activity, number of houses, population density. Input-orientated VRS.	Second-stage regression.	Network and population density has a significant influence on efficiency.
Kirkpatrick, Parker and Zhang (2006)	SFA and DEA	110 public and private water utilities, Africa, 2000.	Dependent/Input: Operating and maintenance expenditure. Independent: Labour price, material price of water distributed, number of water treatment works. Output: Water delivered, hours of piped water per day. Input-orientated VRS.	Descriptive analysis.	No evidence of better performance of private utilities over state-owned utilities. Impact of water technology, transactions costs and regulation on efficiency scores.
Saal and Parker (2006)	MI and SFA	10 public regional water authorities and 29 private statutory water and sewerage companies, England and Wales, 1993-2003.	Inputs: Inputs: fixed physical capital, operating expenditure. Outputs: Water delivered and number of connected properties. Nondiscretionary inputs: population served per kilometre length of mains (density), average pumping head and average quality compliance, dummy for water and sewerage company. Input-orientated CRS and VRS.	Descriptive analysis.	Scope for use of techniques in measuring operational efficiency. Inappropriate to assume water authorities and water and sewerage companies share a common frontier.
da Silva e Souza, Coelho de Faria and Moreira (2007)	SFA	149 public and 15 private companies, Brazil 2002	Dependent: Average costs. Independent: Volume of water produced, prices of capital and labour, average tariff. Explanatory: Private and public utilities, population density, percentage of above ground water sources, regional dummies.	Single-stage regression.	No evidence that private and public utilities differ in estimated efficiency. Significant impact of environmental factors.

Author(s)	Methodology[a]	Sample[b]	Specification[c]	Technique	Findings
García-Valiñas and Muñiz (2007)	DEA	3 water supplying municipalities, Spain, 1985–2000.	Input: Operational expenditures. Output: Volume of water delivered, length of mains, population supplied. Nondiscretionary input: Rainfall. Input-orientated CRS.	Descriptive analysis.	Inclusion of nondiscretionary factors increases observed level of efficiency.
Saal, Parker and Weyman-Jones (2007)	SFA	England and Wales, 10 water and sewerage companies, 1985–2000	Dependent: Water customers, connections with sewerage customers, physical water supply, physical sewerage load; quality adjustment indices (water and sewerage). Independent: Capital stock, current cost operating profits less current cost depreciation, infrastructure renewal expenditures, non-capitalised employment, labour.	Descriptive analysis.	Technical change improved after privatisation but not productivity growth. Excessive size of water supply companies contributed negatively to productivity growth.
Filippini, Hrovatin, and Zoric (2008)	SFA	52 water utilities, Slovenia, 1997–2003	Dependent: Total annual cost. Independent: Prices of labour, capital and materials, water supplied, number of customers, size of service area, treatment dummy, dummies for surface water, groundwater and low water losses.	Descriptive analysis.	Inefficiency estimates depend on econometric specification. Diseconomies of scale in larger utilities.
Picazo-Tadeo, Sáez-Fernández, and González-Gómez. (2008)	DEA	40 Spanish water utilities (with 20 providing sewerage services), 2001.	Inputs: Delivery network, sewer network, labour, operational costs. Outputs: Population served, water delivered, treated sewage. Output-orientated CRS.	Descriptive analysis.	Accounted-for water does not influence ranking of utilities. Quality matters in measuring technical efficiency.
Guder, Kittlaus, Moll, Walter and Zschille (2009)	DEA	373 water utilities, Germany, 2006.	Input: Total revenue. Outputs: Number of water meters, water delivered to households and non-households (industrial and other), network length, population. Nondiscretionary inputs: Length of network, leak ratio, groundwater ratio, elevation differences, dummy for former East Germany. Input-orientated CRS and VRS.	Second stage regression.	Substantial differences in technical inefficiency after inclusion of structural factors. Network density and share of groundwater negatively influence efficiency.
Reznetti and Dupont (2009)	DEA	64 Canadian water utilities, 1996.	Inputs: Labour costs, materials costs, delivery network. Outputs: Water delivered. Nondiscretionary inputs: Extreme temperatures, precipitation, dummy for surface water, population density, elevation, proportion of residential demand, number of dwellings. Input-orientated VRS.	Second-stage regression.	Differences in elevation, population density, and proportion of residential water use private dwelling have significant impact on efficiency.
Byrnes, Crase, Dollery and Villano (2010)	MI	14 Victorian water utilities and 38 NSW water utilities, 2000–04.	Input: Total operating costs. Outputs: Complaints index and total potable water delivered. Nondiscretionary inputs: Proportion of residential consumption, water losses, production density, customer density, large and very large utilities, share of groundwater, filtration and reticulation dummies, dam maintenance, temperature, rain days, rainfall, rainfall intensity, state identifier, yearly dummies.	Second-stage regression.	Water restrictions reduce efficiency and larger utilities characterised by higher efficiency.

Author(s)	Methodology[a]	Sample[b]	Specification[c]	Technique	Findings
Munisamy (2010)	DEA	6 water supply authorities and 11 privatised water companies, Malaysia, 2005.	Inputs: Operating expenditure, network length, volume of non-revenue water. Outputs: Volume of water delivered, number of connections, size of service area. Input-orientated CRS and VRS.	Descriptive analysis.	Scale inefficiencies in (smaller) private sector utilities, technical inefficiencies in public providers.

Notes: (a) DEA – Data Envelopment Analysis, SFA – Stochastic Frontier Analysis, MI – Malmquist Indices, CRS – constant returns-to-scale, VRS – variable returns-to-scale; (b) Single dates are calendar or financial year cross-sections, intervals are time-series; (c) Specification SFA comprises dependent, independent and explanatory variables, DEA and MI is discretionary input(s), discretionary output(s) and nondiscretionary input(s); (d) All SFA studies usually discuss the estimated coefficients, significance and elasticities for the production and cost parameters, as well as the measures of efficiency obtained. Descriptive analysis includes analysis of distributions (mean, standard deviations) and/or analysis of efficiency by groups within sample and correlation between efficiency scores obtained by different techniques. Second-stage regression involved regressing efficiency scores from DEA, MI, or SFA on additional explanatory variables in a separate regression (usually Tobit, probit or logit), single-stage regression refers to a stochastic frontier model where efficiency estimates are estimated simultaneously with the coefficients on the explanatory variables.

Source: *Worthington (2010)*

REFERENCES

Abbott, M. & Cohen, B. (2009), *Productivity and efficiency in the water industry*, Utilities Policy.

Baltagi, B.H. (2005), *Econometric Analysis of Panel Data* Chichester, Wiley.

Baumol, W J, Panzar, J C & Willig, R D (1982), Contestable Markets and the Theory of Industry Structure, London, Harcourt Brace Jovanovich.

Berg, S.V.(2010), *Water Utility Benchmarking: Measurement, Methodologies, and Performance Incentives*, London: IWA Publishing.

Berndt, E.R. & Christensen, L.R. (1973), *The Internal Structure of Functional Relationships: Separability, Substitution, and Aggregation* The Review of Economic Studies, Vol. 40, No. 3 pp. 403-410.

Bogetoft, P. & Otto, L. (2011), *Benchmarking with DEA, SFA, and R* London: Springer.

Bolt, C. (2010), *Introducing the concept of the system operator in the water industry: lessons from the gas and rail sectors in the UK.*

Bottasso, et al. (2011), *Technologies: Evidence from the English water and sewerage utilities* International Journal of Production Economics 130 112–117.

Bottasso & Conti (2009a), *Price cap regulation and the ratchet effect: a generalized index approach* Journal of Productivity Analysis, DOI 10.1007/s11123-009-0140-z.

Bottasso & Conti (2009b), *Scale economies, technology and technical change in the water industry: Evidence from the English water only sector* Regional Science and Urban Economics 39 (2009) 138–147.

Chambers, R.G. (1988), *Applied Production Analysis*. Cambridge: Cambridge University Press.

Coelli, T. & D. Lawrence (Eds) (2006), *Performance measurement and regulation of network utilities* (pp. 297-327).

Coelli, T., Perelman, S., & Romano, E. (1999), *Accounting for environmental influences in stochastic frontier models: with application to international airlines.* Journal of Productivity Analysis, 11(3), 251-273.

Coelli, T, et al. (2003). *A Primer on Efficiency Measurement for Utilities and Transport Regulators.* Washington: The World Bank.

Coelli, T., et al. (2005), *An Introduction to Efficiency and Productivity Analysis*, Springer

Competition Commission (2010), *Bristol Water plc Price Limits Determination*.

Competition Commission (2007), *A report on the completed water merger of South East Water Limited and Mid Kent Water Limited.*

Erbetta, F. and Cave, M. (2006), *Regulation and efficiency incentives: Evidence from the England and Wales water and sewerage industry*, Review of Network Economics, 6(2), 425–452.

Evans, D.S. Heckman, J.J.(1984), *A Test for Subadditivity of the Cost Function with an Application to the Bell System* The American Economic Review*,* Vol. 74, No. 4, pp. 615-623.

Frontier (2010), *RPI-X@20: The future role of benchmarking in regulatory reviews, A final report prepared for Ofgem*, Ofgem.

Greene, W.H. (2008), *The Econometric Approach to Efficiency Analysis, in Fried*, H.O, et al. (Eds) The Measurement of Productive Efficiency and Productivity Growth*,* Oxford: Oxford University Press.

Greene, W.H.. (2005), *Fixed and Random Effects in Stochastic Frontier Models,* Journal of Productivity Analysis, Springer Netherlands.

Gort, M. & Sung (2009), *Competition and productivity growth: The case of the U.S. telephone industry*, Economic Inquiry, Volume 37, Issue 4, pages 678–691, October 1999.

Hackman, S.T. (2008), *Production Economics: Integrating the Microeconomic and Engineering Perspectives*, Berlin: Springer-Verlag.

Kumbhakar, S. & Horncastle, A. (2010), *Improving the econometric precision of regulatory models*, Journal of Regulatory Economics, 38(2), 144-66.

Lawrence, D., & Diewert, E., (2004) *Measuring Output and Productivity in Electricity Networks*, SSHRC Conference on Index Number Theory and the Measurement of Prices and Productivity Vancouver, 30 June – 3 July 2004.

London Economics (2010), *Competition in upstream sewage and sludge markets, A report prepared for Ofwat,* Ofwat.

Maziotis, A. (2010) *Developing Panel Based Methods for Profit and Performance Measurement in the Water and Sewerage Sector in England and Wales* PhD Dissertation, Aston Business School, Aston University, UK.

Maziotis, Saal, & Thanassoulis (2011), *A panel based profit decomposition technique to identify the sources of profit change in the English and Welsh Water and Sewerage Companies*, Aston Centre for Critical Infrastructure and Services working paper, Aston University, UK.

Monopolies and Merger Commissions (1996), *Severn Trent Plc and South West Water Plc – a report on the proposed merger*.

Ofwat/ UKWIR (2007), *Review of the Approach to Efficiency Assessment in the Regulation of the UK Water Industry*, UKWIR.

Ofgem (2011a), *Decision on strategy for the next transmission price control, RIIO-T1 Tools for cost assessment,* 31 March 2011, Ofgem.

Ofgem (2011b), *Decision on strategy for the next transmission price control – RIIO-T1*, 31 March 2011, Ofgem.

Ofwat (2009), Cost base feedback report: December 2009.

Ofwat (2011a), *Future price limits – a preliminary model: informal consultation*, April 2011.

Ofwat (2011b), A hypothetical model for upstream water markets in England and Wales – a technical paper.

Oxera (in association with Professor Subal Kumbhakar), *Bristol Water's Efficiency,* independent submission to the Competition Commission, April 2010.

Reckon (2008), *PR09 Scope for efficiency studies, A report prepared for Ofwat.*

Saal, et al (2011), *Scale and scope economies and the efficient configuration of the water industry: a survey of the literature* Aston Centre for Critical Infrastructure and Services Working Paper, Aston University, UK.

Saal, D S, Arocena, P & A Maziotis (2010), *Measuring Economies of Horizontal and Vertical Integration in the English and Welsh Water and Sewerage Industry* Aston Centre for Critical Infrastructure and Services working paper , Aston University, UK.

Saal D. & Parker D. (2000), *The impact of privatization and regulation on the water and sewerage industry in England and Wales: A translog cost function model*, *Managerial and Decision Economics*, 21, 253-268.

Saal, D., Parker, D. & Weyman-Jones, T. (2007), *Determining the contribution of technical efficiency, and scale change to productivity growth in the privatized English and Welsh water and sewerage industry:* 1985-2000. Journal of Productivity Analysis, 28 (1), 127-139.

Simar, L. Wilson, P. (2008). *Statistical Inference in Non-parametric Frontier Models: Recent Developments and Perspectives*, in Fried, H.O, et al. (Eds),The Measurement of Productive Efficiency and Productivity Growth, Oxford: Oxford University Press.

Stone & Webster (2004a), *Investigation into evidence for economies of scale in the water and sewerage industry in England and Wales: Final Report, prepared for and published by Ofwat.*

Stone & Webster (2004b) *An investigation into opex productivity trends and causes in the water industry in England & Wales* – 1992-93 to 2002-03, prepared for and published by Ofwat.

Thanassoulis, E. (2001) *Introduction to the Theory and Application of Data Envelopment Analysis*, Kluwer Academic Publishers.

Thanassoulis, et al. (2008). *Data Envelopment Analysis: The Mathematical Programming Approach to Efficiency Analysis,* Fried et al. (Eds),The Measurement of Productive Efficiency and Productivity Growth, Oxford: Oxford University Press.

UKWIR (2008), *Application of Time Series Analysis to Relative Efficiency Assessment*, UKWIR.

Weyman-Jones, et al. (2006) *Efficiency Analysis for Incentive Regulation* (In Coelli, T. & D. Lawrence (Eds), Performance measurement and regulation of network utilities.

Worthington, A. (2010) *A review of frontier approaches to efficiency and productivity measurement in urban water utilities*, Griffith University, Australia.