OFWAT

COST ASSESSMENT

22nd January 2013

Prepared by:

Cambridge Economic Policy Associates Ltd

CONTENTS

IMPORTANT NOTICE

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

GLOSSARY

Term	Definition
ASHE	Annual Survey of Hours and Earnings prepared by the Office of National Statistics
Baseline	A cost value, derived from the model forecasts/company business plan forecasts, which is used in the menu
BCIS	Building Cost Information Service
Capex	Capital expenditure
Cobb-Douglas model	The Cobb-Douglas (or log-linear) model transforms the variables into logarithms prior to estimation. This model is deemed superior to a linear model in the cost modelling literature as it does not require marginal costs to be constant as in the linear model. Even so, the Cobb-Douglas model is in itself restrictive because, inter alia, it assumes that the extent of returns to scale is the same irrespective of firm size. Compare with translog model.
Corrected OLS (COLS)	Please refer to ordinary least squares (OLS). 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.
Correlation (coefficient)	A correlation coefficient is the measure of interdependence between two variables. The value ranges from -1 to 1, with -1 indicating a perfect negative correlation and 1 indicating a perfect positive correlation. Zero indicates the absence of correlation between the variables.
COPI	Construction price index prepared by BCIS
Corridor	The range calculated by using the model parameters, against which company cost forecasts are evaluated
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.
FPL	Future Price Limits
Hausman test	This test provides information on whether the fixed or random effects treatment is most appropriate. A high value of the statistic (which represents a rejection of the null hypothesis) indicates that the fixed effects model is preferred to the random effects model. Otherwise the random effects treatment is preferred.
I&C	Industrial and commercial customers
IRC	Infrastructure renewals charge (annual allowance)
IRE	Infrastructure renewal expenditure (actual)
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.

Term	Definition
MNI	Maintenance of non-infrastructure expenditure (actual)
Multicollinearity	When two or more explanatory variables are highly correlated with each other. This may mean that in a multiple regression model the coefficients on the correlated variables may not be valid, but the overall predictive power of the model is not reduced (only the ability to use the coefficients individually).
Opex	Operating expenditure
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.
Pooled OLS	The pooled corrected ordinary least squares model treats the data as if it were a cross-section – that is, e.g. 90 firms, rather than a panel of 10 firms over nine years. This approach does not therefore recognise the panel structure of the data, and can be tested against the panel model variants. It is however a simple model that is used by economic regulators in particular.
Pooled Stochastic Frontier Analysis (SFA) model	This is a maximum likelihood estimation model that is the same as COLS except that a one-side error term is included to permit the existence of inefficiency (with the error term decomposed into its noise and inefficiency components). This approach requires distributional assumptions about those components.
PR14	Price Review 2014
Time invariant efficiency model: Fixed Effects (FE)	This is the standard fixed effects model used in the panel data literature, except that in this case the fixed effects terms are given an inefficiency interpretation.
Time invariant efficiency model: Random Effects (RE)	This is the standard random effects model used in the panel data literature.
Time-invariant SFA model	This is a maximum likelihood model and an extension of the random effects model but now with distributional assumptions imposed and with estimation proceeding via MLE, not generalised least squares (GLS), as in the standard panel data random effects model. See Pitt and Lee (1981).
Time varying SFA model	This is a maximum likelihood model that extends the model above to permit efficiency to vary over time but in a restricted way, since the direction of efficiency change over time must be the same for all firms (and thus rankings cannot change). See Battese and Coelli (1992)
STW	Sewage treatment works
STW Band	Ofwat defines six bands of STW depending on their size. Band 1 relates to the smallest STWs and Band 6 the largest. We have further split Band 6 into two bands, with the new Band 7 defining the very largest STW
Totex	Total expenditure (opex + capex)

Term	Definition
Translog model	The translog model is one of the so called flexible functional forms and is used routinely in the academic literature. In the current context one of its particular advantages is that it allows the degree of returns to scale to vary with firm size. The Cobb-Douglas is nested within the translog so it is possible to test the Cobb-Douglas restriction.
Triangulation	The use of multiple methodologies and the numbers from them (averages, max, min etc) to come up with a single value for cost assessment
UKWIR	UK Water Industry Research
WaSC	Water and sewerage company
WoC	Water only company
WTW	Water treatment works

EXECUTIVE SUMMARY

Cambridge Economic Policy Associates Ltd (CEPA) was requested by Ofwat to provide support on the question of cost assessment for PR14 and beyond. This built on work undertaken by:

- CEPA on Incentives and Menus for Ofwat, published in August 2012 alongside the Ofwat wholesale incentives consultation document; and
- Reckon for UKWIR on cost assessment and incentives which was completed in September 2012.

Specifically, Ofwat asked CEPA to determine the viability of total cost or totex assessment for PR14 and beyond. In addition, if it was found that a totex assessment was not possible using integrated models, CEPA was asked to advise on alternative approaches that could be employed.

Methodology

We approached the viability assessment by first testing whether integrated totex cost models were viable for water and sewerage. If we could not establish viability at the integrated totex level we tested the viability of cost modelling to support totex cost assessment at a more granular level of costs (i.e. opex, opex plus base capex, and enhancement capex) and at the sub-company level (e.g. sewage treatment works).

We have defined our cost models as being a function of three broad types of explanatory variable. The three broad types of variable that we have considered are:

- scale variables – these define the broad overall size of the business and capture elements like the length of the network, the density of the population using the network etc;
- company characteristic variables – these capture specific aspects of the company that are outside the control of management. Examples of the types of variable that were included here are pumping head and regional wage and construction prices; and
- activity and quality variables – these capture differences in possible cost drivers like metering, quality of service, new connections etc.

The cost driver data were sourced from the June Returns, with additional data being provided by Ofwat for average wages and the COPI.

Our assessment of viability has been based on at least one of our models meeting each of the following: (i) cost drivers which are statistically significant; (ii) the form and 'direction' of coefficients meets expectations (i.e. realistic cost function represented); and (iii) efficiency estimates that seem plausible given Ofwat's previous modelling and determinations.

Where we were unable to find viable totex models, even at the more granular level, we then considered other approaches to support the derivation of totex corridors and baselines. These have included unit cost based approaches as well as opex econometric models.

Conclusions

We found that:

- integrated totex cost modelling at the overall water wholesale level is viable;

- water wholesale opex with base capex models are also viable;
- water wholesale enhancement capex models were not robust;
- at the overall wholesale level, integrated totex cost assessment is not a viable approach for sewerage; and
- at the sewerage network level, cost assessment for base costs (opex and capital maintenance) is viable.

We considered what alternative approaches might work for sewerage (noting that separate approaches to cost assessment for sludge and private sewers activities are already being developed by Ofwat). A mix of econometric opex models and unit cost based approaches appear viable to support totex cost assessment, and can offer improvements on Ofwat's previous approaches as panel data allows the use of trend analysis and efficiency challenge.

These results are summarised in Table E1 below.

Table E1: Viable cost assessment approaches for Wholesale Water and Sewerage totex expenditures

	Opex	**Base capex**	**Enhancement capex**
Water wholesale (integrated totex models)	Econometric modelling		
Water wholesale (by expenditure type)	Econometric modelling		Cost base/ unit costs
Sewerage network	Econometric modelling		Cost base/ unit cost[1]

Consequently, we tested alternative sub-company approaches to sewerage treatment works and concluded:

- viable opex models exist for size bands 1-5, the smaller treatment works, to be considered at a company level;
- an improved approach to opex unit costs is viable for bands 1-3;
- bands 4-7, the medium and large treatment works, can be viably modelled for opex at a band level; and
- bands 6 and 7, the largest treatment works, can be viably modelled for opex at a treatment works level.

These results are summarised in Table E2 below.

[1] In the main consultation document Ofwat refers to these are models similar to those previously used by Ofwat.

Table E2: Viable cost assessment approaches to Sewage Treatment opex

Opex	Econometric model			Unit costs
	Company level	Company (by band)	Sub-company (by works)	Company (by band)
Band 1	✓	×	×	✓
Band 2		×	×	✓
Band 3		×	×	✓
Band 4		✓	×	×
Band 5		✓	×	×
Band 6	×	✓	✓	×
Band 7		✓	✓	×

So, in summary, we believe that integrated totex models are:

- viable for wholesale water; but
- have limited application to sewerage except for the network.

In part the limited application to sewerage may be driven by data availability since enhancement capex cannot be easily split between different activities.

Alternative approaches are, however, available and can be used to deliver what we believe to be robust results. These include a mix of unit cost and econometric approaches, along the lines of the approaches to cost assessment which Ofwat has previously adopted in this area. While boundary issues clearly arise these are fewer than used to exist with Ofwat's previous cost assessment approach.

1. INTRODUCTION

Cambridge Economic Policy Associates Ltd (CEPA) was requested by Ofwat to provide support on the question of cost assessment for PR14 and beyond. This built on work undertaken by:

- CEPA on Incentives and Menus for Ofwat, published in August 2012 alongside the Ofwat wholesale incentives consultation document; and
- Reckon for UKWIR on cost assessment and incentives which was completed in September 2012.[2]

Specifically, Ofwat asked CEPA to determine the viability of a total expenditure/cost approach to cost assessment. Historically, Ofwat has modelled opex and capex, and base and enhancement separately for both water and sewerage. It should be noted that CEPA were not asked to provide final models for cost assessment at PR14 – further work will be needed to determine such final models once the overall approach(s) to cost assessment have been finalised, which Ofwat has planned throughout 2013.

With this in mind, we do not report specific models in this report but rather discuss:

- the approach to determining viability of models;
- the possible uses of cost assessment at PR14 and beyond and the implications that this has for model choice;
- our approach to modelling; and
- the overall results of the modelling.

At the end of the report we also suggest some possible further considerations for Ofwat and the industry with respect to cost assessment. In part this is linked to the implications of potentially having to use different approaches to assess costs for different elements of the industry or different costs within an element (say enhancement as opposed to base expenditure).

Our work has been supported by significant input from Ofwat, the companies that were members of the UKWIR project steering group for cost assessment and Reckon themselves. We would like to thank everyone for their support.

Dr. Andrew Smith, an expert in efficiency analysis and its uses in regulated industries from the University of Leeds, has provided ongoing support to the project team. This support has ensured a high level of academic rigor has been applied when developing the models and testing their validity.

1.1. Structure of the report

The general coverage of the report was noted above. The specific structure is as follows:

- Section 2 sets out our assessment criteria;
- Section 3 sets out our approach to modelling;

[2] More information is available at the UKWIR website, the project is called "A total expenditure approach to comparative efficiency", http://www.ukwir.org/content/default.asp?PageId=74343.

- Section 4 provides some overall model outputs;
- Section 5 addresses the approach to determining a possible cost corridor;
- Section 6 considers the question of determining a baseline cost; and
- Section 7 sets out lessons from this project.

The remainder of this section discusses the possible uses of cost assessment at PR14.

1.2. Cost assessment and its uses

Before starting to consider the detail of the modelling that we have undertaken, it is worth considering the possible uses that Ofwat might make of cost assessment at PR14. These have been touched on in other reports, including our August Incentives and Menus report, but are briefly recapped below.

1.2.1. Corridors

The first possible use of cost assessment is linked to the triggers or "flags" that Ofwat might consider when reviewing a company's business plan. We understand that Ofwat expect to focus on a more 'light-handed' approach to PR14. As such, Ofwat intends to only undertake greater scrutiny of those business plans which raise concerns about risks to achieving its statutory duties. How these business plans are identified is still under consideration, in the context of the forthcoming consultations by Ofwat, but it is clear that the revenue requirement proposed by the company is one of the key elements of the business plan and as such could be a trigger for review.

How could the revenue requirement be assessed? One of the core inputs to any revenue requirement assessment will be cost assessment. A tool based on the cost assessment that could be used by Ofwat is that of a possible cost corridor, discussed in detail later in this report. If an expenditure proposal (which is the basis of the revenue requirement) sits within the cost corridor then it has passed this cost assessment test for the relevant expenditure.[3] This does not mean that there will not be any further review of the company's revenue requirement. If other tests or triggers are failed then Ofwat may need to consider the costs in the business plan from other perspectives. Consequently the design of the cost assessment corridor depends in part on the other elements of Ofwat's business plan assessment.

1.2.2. Setting a baseline

A second use is that of setting a cost baseline. While the baseline question was discussed in some detail in the Incentives and Menus report, the question of setting a baseline goes beyond the question of the menu. Even if no menu is applied, Ofwat will need to have a baseline value on which it can determine the revenue requirement, and cost assessment is an important part of this.

[3] As we focus on modelling expenditure rather than costs the revenue requirement will be a function of both the proposed expenditure and other decisions. This is because expenditure considers capex while the revenue requirement incorporates the return on and of the expenditure included in the RCV. This is discussed in section 4 of our July 2012 report for Ofwat on Incentives and Menus, published as part of the August consultation.

One possible link to the cost corridor question that has been proposed is that if a company's revenue requirement as set out in its business plan is within the cost corridor (and no other triggers lead to a review) then Ofwat could accept the company's proposal as the baseline. This question is also discussed later in the report.

2. ASSESSMENT CRITERIA

2.1. Background

In order to assess whether the possible approaches to total expenditure modelling are viable for use at PR14 for cost assessment, we have developed a set of assessment criteria. Assessment criteria may be useful in judging cost assessment approaches, but an important consideration is, of course, that some cost assessment approaches may meet different criteria to a greater or lesser extent. As such, it will be important to assess the approaches taking a balanced view of the overall alignment with Ofwat's regulatory objectives. The criteria we have identified are consistent with the wider Future Price Limits (FPL) criteria and Ofwat's wider criteria and are deemed to be fit for purpose for this report.

2.2. Assessment criteria

Any cost assessment approach, especially one linked to total expenditure, would need to fit with Ofwat's FPL principles.[4] These principles are as follows:

Box 2.1: FPL principles

- Targeted price controls
- Proportionate price setting
- Effective incentives
- Ownership, accountability and innovation
- Flexibility and responsiveness
- Transparency and predictability

In the *From principles to price setting – next steps* document,[5] Ofwat chose the following seven assessment criteria to help chose between options:

1. Statutory duties: This relates to Ofwat's statutory duties, which include ensuring a fair deal for consumers (bill impact), making sure companies remain able to finance themselves (financeability impact) and making monopolies improve (promote innovation).
2. Legislation & wider government policy: Remaining consistent with government policy and with a water sector White Paper promoting innovation.
3. Impact: Looking at the benefits and costs of each menu bundle.
4. Inherent risk: Seeing how risk is allocated.
5. Critical path/ dependency: Analysing how the imposition of one regulatory regime may affect future price controls, what is feasible for PR14, and the role of customer engagement.

[4] Ofwat (2012) 'Future Price Limits – A Summary,' May 2012. Available at: http://www.ofwat.gov.uk/future/monopolies/fpl/pap_pos201205fplsummary.pdf

[5] Ofwat (2012) 'From principles to price setting – next steps,' May 2012. Available at: http://www.ofwat.gov.uk/future/monopolies/fpl/pap_pos201205fplsteps.pdf

6. Practical issues of implementation: Several specific issues will occur in trying to implement the preferred option so that problems at the implementation stage may affect the choice of cost assessment approach.
7. Consistency with FPL principles: Ensuring consistency with the six FPL principles listed in Box 2.1.

Looking at the assessment criteria set out above, there is an inevitable degree of overlap between categories. This is especially the case when looking at the third objective (estimating impact) and each of the other criteria. The ability to assess different cost assessment approaches is dependent upon having a clear and concise set of assessment criteria in place. In order to make our assessment of the very long list of incentive options less unwieldy, we have grouped criteria, based on the criteria described above. These are set out in Table 2.1 below.

Table 2.1: Assessment criteria

Criteria	**Description**
Targeted price control	As per the FPL principles, this assessment criterion refers to reducing the complexity of any price-setting tool and having this under a single regulatory framework. The approach will also be focussed on the areas in which it is deemed most necessary. This could also cover the regulatory burden imposed by an approach, although there will be overlap with implementation issues in this respect.
Effective incentives	This point refers to giving effective truth-telling properties and eliminating any unequal treatment of cost that may give rise to a potential capex bias. This should also refer to not giving perverse incentives where companies benefit from inflating their forecasts in one period to benefit in subsequent price controls. Finally, this also needs to relate to the provision of appropriate incentives for efficient delivery of the desired outcomes. This ensures compatibility with Ofwat's proposed menu and outcomes-based approaches.
Statutory duties and impact	This refers to the potential impact of a new approach to both financeability of companies and the impact on bills and value for money for consumers. These are both statutory duties for Ofwat and consequently key assessment criteria. Bill and value for money impacts will include: (i) any step change in bills into PR14; (ii) the degree of annual variation caused by the mechanism; and (iii) the issue of intergenerational equity if the cost burden falls more heavily on certain sets of consumers i.e. in certain time periods. In addition, this would consider broader regulatory objectives such as economic efficiency.
Implementation	While certain approaches may possess theoretical appeal, there are a range of implementation issues that will determine the success of any regulatory framework including information requirements and regulatory burden. It is also important to consider the potential for gaming or strategic interaction. Also there is the issue of the clarity of the regime and the cost assessment approach employed. The regime must be understood by companies in order for them to respond to the right incentives.

Our judgement is that these criteria cover the main points in how best to assess the relative merits of a cost assessment approach. However, the breadth of the areas means that the table is not fully inclusive of all criteria and, similarly, it may be beneficial to look at narrower questions within the above criteria.

It will also be necessary to analyse implementation issues in more detail, and specifically to take account of possible splits of the value chain. Some options will score better in terms of implementation if multiple splits are envisaged. It may also be the case that some forms of cost assessment work at one level of aggregation but not another. This applies in particular for example for options which require data to be available at the disaggregated value-chain level.

Linked to this last point, and one of the key drivers of the reforms that Ofwat is considering, is the issue of effective incentives and the problems that arise when boundaries exist. When we are considering cost assessments at more granular levels we will need to be aware of the incentive implications of this.

3. COST ASSESSMENT MODELLING

3.1. Overall approach

The specific question posed to us, whether a total expenditure assessment approach is viable raises several key modelling questions. These are discussed below. First, there is the question of the overall approach. We consider this question from two different dimensions:

- First – is the approach adopted an econometric one? And
- Second – is a direct estimate of the costs being made?

Each is discussed in turn.

3.1.1. Econometric and non-econometric

There are many ways in which cost assessment can be undertaken for planned expenditure. These include:

- simple unit cost comparisons;
- engineering based model company approaches; and
- detailed econometric model based approaches (ranging from least squares type regression modelling to frontier based approaches of data envelope analysis etc).

Our focus in this report is primarily on the question of whether econometric total expenditure/ cost assessment approaches are viable. That is not to dismiss the other approaches and we return to one of these, unit cost measures, later in the report.[6]

3.1.2. Direct estimation vs roll-forward

The question of what is being determined is important when thinking about the use of cost assessment. Ofwat's traditional approach, based around separate opex and capex assessments, could be considered as a way of determining the relative efficiency of companies (in bands for opex). The opex relative efficiency was then applied to a starting year that was rolled forward with an assumed speed of catch-up over the life of the price control, while the relative efficiency for capital maintenance was applied to the forecast expenditure for each year. This is a quite standard approach to cost assessment and the creation of performance incentives and is suited to a separate consideration of base and enhancement expenditure. It is less appropriate when thinking about totex.

An alternative more direct method, better suited to totex cost assessment, would be to use the cost assessment models to determine the actual level of required revenue for a year, or the whole

[6] Engineering based approaches have not been viewed as viable approaches in the UK. They are briefly discussed in our Incentives and Menus August 2012 report. Ofwat's cost base approach, using sample standardised investment project costs as a way of assessing relative efficiency, is an example of a (relatively) simple unit cost approach being used to set efficiency challenges. This approach, however, was not a full cost assessment strategy and, in principle, cannot be used for this purpose. If the cost base was to be used as the main basis to set the overall revenue allowances, a more comprehensive approach would be needed.

price control period. A separate cost performance incentive, through a mechanism like the proposed menu, would then be implemented alongside this.

The willingness to place a greater emphasis on the cost assessment models, which the direct estimation approach does, is driven by several considerations. These include:

- the ability to have more comprehensive cost assessment models going forward;
- the robustness created by having a range of acceptable estimated costs within a cost corridor (so allowing for the possible lack of precision in individual models); and
- the ability of the company to address any lack of precision through the menu or whatever performance incentive approach is adopted by Ofwat (although the willingness of companies to accept this will be linked to the overall menu design).

We have taken as our default position the proposal that direct estimation is the preferred position.

3.1.3. Form of dependent (left hand side) variable

When considering the modelling of total cost/expenditure, there are two basic approaches to defining the left hand side variable – the "costs" that are being explained. Specifically:

- total expenditure (totex), i.e. opex and capex; or
- total costs, i.e. opex plus the return on and of capital.

In the latter approach much of the capex is captured through the impact that it has on the RCV and subsequently the return on and of capital while in the former capex is captured directly. There are advantages and disadvantages with both approaches, as set out in table 3.1 below.

Table 3.1: Strengths and weaknesses of different cost measures

Measure	Advantages	Disadvantages
Totex	Reflects the expected expenditure over the period Simple to understand Cost performance incentives are focused on expenditure	Does not directly reflect the allowed revenue that is the basis of the regulatory settlement When considering quality/ outcomes the stock of capital is likely to also have an influence but is not captured: however, the approach is in part reliant on alternative regulatory safeguards such as the serviceability requirements Ofwat has already put in place for the current control period.
Total costs	Mirrors the allowed revenue and so makes it simple to understand the implications of the cost assessment for prices May capture any impact that the stock of capital has rather than just the flow has on outputs/outcomes	Changes in the capital stock may not reflect the impact that capex is actually having on costs and outcomes Great care is needed to ensure that the measures of return on and of capital are appropriate (could lead to gaming if companies can build arguments for differences in their depreciation rates etc)

Measure	Advantages	Disadvantages
		More complex than expenditure measures

There are regulatory examples of both approaches being used:

- Ofgem has modelled totex for RIIO-GD1 (see Annex 1) and is proposing to use the same approach for RIIO-ED1;
- ORR has employed totex modelling for Network Rail; and
- NMa (the Dutch regulator) has used total cost modelling for energy distribution and transmission networks in Holland.

Ofwat has proposed to use totex and consequently we consider this approach in the remainder of the report.

3.2. Econometric approaches

As noted above, we have primarily focused on econometric approaches when considering cost assessment approaches. Econometric approaches have generally been preferred by Ofwat and other UK regulators in previous approaches to cost assessment. Econometric approaches have a wide range of attributes that can make them preferable to alternative approaches, for instance:

- they allow the selection of specific functional forms;
- they are applicable to a small number of observations (although greater numbers are preferred) right through to a very large number; and
- they allow for comparative analysis.

This section describes the approach to modelling that we have adopted, the data employed and the types of models considered.

3.2.1. Approach

Our approach to modelling can be considered in two ways:

- the basic elements of any model we consider; and
- the way that we have moved from general to specific models.

Elements of the model

We have defined our cost models on the basis of totex being a function of three broad types of explanatory variable. The three broad types of variable that we have then considered are:

- scale variables – these define the broad overall size of the business and capture elements like the length of the network, the density of the population using the network etc;
- company characteristic variables – these capture specific aspects of the company that are outside the control of management. Examples of the types of variable that were included here are pumping head and regional wage and construction prices; and

- activity and quality variables – these capture differences in possible cost drivers like metering, quality of service, new connections etc.

Using Ofwat's existing information database, primarily from the June Returns, selections of variables were collected and fit into these three categories. This is discussed further in the sub-section on data below.

Refining the model

In each initial model we included all available explanatory variables under the three broad headings. We then investigated whether the models were improved, or at least no worse, if some of the variables were excluded. Choosing variables to exclude was based on:

- statistical significance – those that did not seem to be significant were tested for exclusion;
- sector significance – is the variable one that *a priori* is expected to be a significant explanatory variable; and
- appropriateness of the result – is the sign and impact of the variable what would *a priori* be expected?

With respect to the last exclusion factor, considering the robustness of the explanation for any variable included was important. The latter two exclusion factors are particularly important as focusing only on the statistical significance of variables may result in a mis-specified model. High correlation between variables, which are expected to be significant between cost drivers, may cause the effect of one of them to be picked up by the other variable, thus making the former variable appear less significant.

Through this process we were able to reduce the number of explanatory variables while retaining as significant a model as possible – we refer to these as "refined" models. This was important given the limitations created through the number of observations that were available.

Through-out this work, as noted in the Introduction, Dr Andrew Smith has provided advice to the team. This was especially important when considering the appropriate refinement processes to follow. He has also reviewed the outputs and provided quality assurance for this process.

3.2.2. Data

The starting point for our analysis was the dataset that Ofwat had provided Reckon when it commenced its earlier work for UKWIR.

Our basic dataset covered a period of nine years, namely 2001/02 to 2010/11. This was supplemented with some additional variables, whose possible significance arose during discussion with stakeholders, with much more limited coverage, especially the regional cost adjustments.

With respect to the number of companies covered we used:

- 18 water companies; and
- 10 sewerage companies.

The 18 water companies reflect the mergers that have occurred, especially among water-only companies. Where necessary we have sought to create consistent data series for the combined company although in some cases this required assumptions about the way in which costs should be combined. For example, by simply adding the costs of the two pre-merged companies this means that any economies of scale or scope are not captured until the merger actually occurs. We do not consider this to be a significant limitation to the models because the economies were not realised before 2010-2011.

We also, following discussion with Ofwat, made adjustments to the basic dataset so that we could have an estimate of:

- wholesale water costs rather than total service costs (this entailed making an assumption about the split between retail and wholesale costs prior to the implementation of the accounting separation); and
- base costs (opex and capex) separate to enhancement costs (capex) – this required us to make an assumption that all opex was base while in practice a small proportion of it was actually enhancement. However, we do not feel that this was sufficient to materially affect the results.

Obviously the need to make these sorts of assumptions should be borne in mind when using the resulting cost estimates. However, for the cost assessment to be meaningful for PR14 and beyond these adjustments needed to be made and it will be important to improve upon them where possible when more data becomes available in the process.

Some other data limitations arose in sewerage linked to the issues around sludge and this led to the decision to exclude sludge from our work. Certain cost categories, especially around capex, did not provide a sufficiently granular definition to identify sludge or other segmental information.

The variables considered for water and sewerage under the three broad categories outlined earlier in the section are detailed in the following two tables.

Table 3.2: Range of explanatory variables in water models

Type	**Variable**	Definition/ Transformation	Rationale
Core	Length of mains	Total length of mains at year end	Network scale variable and overall business size proxy (WaSC vs. WoC)
	Property density	Number of connected properties/ length of main	Rural vs. urban divide and economies of density indicator
	Usage	(Potable + non-potable water delivered)/ length of main	Network and resource usage and possible proxy for domestic vs. I&C[7] usage - results similar when normalized by population
	Time trend	Year dummy	Takes into account that the data is for 18 companies over five years and shows the change in costs over the years all other

[7] Industrial and commercial

Type	Variable	Definition/ Transformation	Rationale
			things being equal
Input prices	Average regional wage	Provided by Ofwat in nominal terms for April 2008, 2009, and 2011. The data is based on ONS' ASHE surveys by county and allocates the company's service areas to counties. The wages figure is the average gross weekly salary based on the number of jobs in the company area. The data is transformed to real terms using RPI and 2010 is interpolated as the average between 2009 and 2011. The 2007 estimate is based on an extrapolation using % change for England and Wales.	Input price, one of the main opex and capex drivers; the use of these regional indices does not easily deal with the fact that where companies use contractors they may be brought in from other regions and thus have different underlying input prices. Where contractors are providing capex services the adjustment is made through the COPI estimate rather than the average regional wage
	Regional COPI	Provided by Ofwat in real terms for 2002-2010; extrapolated for 2011 as equal to 2010. The variable uses the construction price index from BCIS and allocates the BCIS areas to the companies based on population numbers from the 2001 census. The index was adjusted by the population proportion served within each area. The index is originally in real terms.	Input price, one of the main capex drivers.[8]
Network characteristics	Population density (occupancy)	Population connected /number of properties connected at year end	Approximates average consumer size (domestic vs. I&C) and can be used to take some of the variation away from usage
	Proportion of metered properties	(Metered billed households with external meters + metered billed households without external meters + metered billed non-households)/ number of properties connected at year end	Metered customers are assumed to have lower per capita consumption than non-metered customers, thus leading to lower pumping and volume related costs; this variable also captures the wholesale costs related to metering such as installation and replacement; during the period covered, some companies entered the replacement cycle and others had significant increases in meter penetration, which would lead to a positive correlation

[8] We understand from Ofwat that the COPI may capture the differences across companies within a year, however it is less comparable across years as the sample within regions is changing.

Type	Variable	Definition/ Transformation	Rationale
			between proportion of metered properties and totex; it is not clear which factor would be stronger
	Proportion of usage by metered household properties	Water delivered to billed metered households/(potable + non-potable water delivered)	In order to estimate the model, one proportion has to be omitted. The omitted variable is non-metered properties and the coefficients on the included variables should be interpreted relative to the one excluded. If the coefficient sign is positive, then metered household properties have higher costs than non-metered properties
	Proportion of usage by metered non-household properties	Water delivered to billed metered non-households/(potable + non-potable water delivered)	The omitted variable is non-metered properties and the coefficients on the included variables should be interpreted relative to that. Proxy for proportion of I&C customers
Treatment and sources characteristics	Sources	Total number of sources/ distribution input	There are economies of scale in the resource and raw water distribution part of the business
	Pumping head	Pumping head x distribution input	Energy proxy: the higher the pumping head and the lift over which water needs to be pumped, the higher the energy usage – used in old Ofwat opex power model
	Proportion of water input from river abstractions	Proportion of water input from river abstractions	Proxy for water treatment works (WTW) complexity; Boreholes are omitted and considering that boreholes water is generally the cheapest type of source to treat, expect signs to be positive
	Proportion of water input from reservoirs	Proportion of water input from reservoirs	Same as above
Activity	Proportion of new meters	(selective + optant meters installed)/ (Metered billed households with external meters + metered billed households without external meters + metered billed non-households)	Enhancement activity
	Proportion of new mains	New mains/Total length of mains at year end	Enhancement activity
Quality	Serviceability	1 if stable or above, 0 if below	Quality measure: if a company changes from 0 to 1 (thus

Type	Variable	Definition/ Transformation	Rationale
	level	stable[9]	increasing quality of base service), it has done so by capex spend
	Properties below reference pressure level	Properties below reference pressure level/total properties connected	Quality measure: the lower the proportion of properties with inadequate water pressure, the higher the costs because companies have spent or are spending money to improve quality but relationship is unclear in the models
	Leakage	Leakage volume/distribution input	Quality measure: the lower the leakage, the higher the costs because companies have spent money to reduce it; however, companies with a lot of leakage will have to spend more to deal with it – does not always work as quality variable
	Properties affected by unplanned interruptions > 3 hrs	Properties affected by planned interruptions > 3 hrs/ total properties connected	Service quality measure: the more interruptions, the lower the quality; thus if interruptions decrease, this might be associated with service enhancement and thus higher costs, particularly because these interruptions are unplanned
	Properties affected by planned interruptions > 3 hrs	Properties affected by unplanned interruptions/ total properties connected	Service quality measure: the more interruptions, the lower the quality; thus if interruptions decrease, this might be associated with service enhancement and thus higher costs; planned interruptions however may be correlated with maintenance works and may result in positive sign

More granular levels of cost information were considered for sewerage as:

- the smaller number of companies meant fewer observations available to support the models which could be overcome both by extending the period considered (but this suffers from issues around stability of management approach etc) or by considering more granular costs which effectively increased the number of observations; and

[9] This is an aggregate score of the water infra and water non-infra serviceability level. Ofwat calculated those based on six factors for infra and five factors for non infra and give companies a score of Improving, Stable, Marginal, Deteriorating or Uncertain every year. We have given companies a 1 if stable or improving and a 0 if anything else. If at least one of the scores (infra and non-infra) for a company in a particular year is 0, the overall score is 0. If both are 1, the score is 1. We have done the same for sewerage. For merged water companies... More info on serviceability levels can be found here: http://www.ofwat.gov.uk/publications/pricereviewletters/ltr_pr0938_appendix2.pdf

- sewerage cost assessment has traditionally proven difficult to model.

Table 3.3: Range of explanatory variables in sewerage models

Type	Variable	Definition/Transformation	Rationale
Core	Length of sewers	Total length of sewers at year end	Network scale variable
	Density	(Household + non household properties connected)/ length of sewers	Rural vs. urban divide and another economies of density indicator
	Usage	Total load entering system/ population connected	Network usage and possible proxy for domestic vs. I&C usage
	Time trend	Year dummy	Takes into account that the data is for 10 companies over nine years and shows the change in costs over the years, all other things being equal
Input prices	Average regional wage	Provided by Ofwat in nominal terms for April 2008, 2009, and 2011. The data is based on ONS' ASHE surveys by county and allocates the company's service areas to counties. The wages figure is the average gross weekly salary based on the number of jobs in the company area. The data is transformed to real terms using RPI and 2010 is interpolated as the average between 2009 and 2011. The 2007 estimate is based on an extrapolation using % change for England and Wales.	Input price is one of the main opex and capex drivers; assumption is that there is little outsourced outside the region of the company's operation
	Regional COPI	Provided by Ofwat in real terms for 2002-2010; extrapolated for 2011 as equal to 2010. The variable uses the construction price index from BCIS and allocates the BCIS areas to the companies based on population numbers from the 2001 census. The index was adjusted by the population proportion served within each area. The index is originally in real terms.	Input price is one of the main capex drivers
Network characteristics	Proportion of metered properties	(Metered billed households + metered billed non households)/ number of properties connected at year end	Metered customers are supposed to have lower per capita usage than non-metered ones, leading to lower costs. Suggested by Ofwat as I&C proxy

Type	Variable	Definition/Transformation	Rationale
Quality	Internal floods	Internal flooding incidents/ (household properties connected + non household properties connected)	Quality measure: the higher the number of floods, the lower the quality, the lower the cost; as flood incidents are also a source of opex, variable may not work as a quality measure but a maintenance driver
Treatment	Load	Total load in relevant category in kg BOD_5[10]/day	Size/scale variable and main cost driver
	Proportion of load in treatment works size bands 1-3	(Load in bands 1-3)/total load for bands 1-5	This variable should be interpreted in reference to proportion of load in size band 5. Since Bands 1-3 tend to be more expensive than 5 in terms of unit costs, it is expected that a higher proportion of 1-3 load would lead to higher costs.
	Proportion of load in treatment works size band 4	Load in band 4/total load for bands 1-5	Similarly to the above, band 4 is expected to be slightly more expensive than band 5, thus leading to the expectations of a positive coefficient.
	Number of works in treatment works size bands 1-3	Number of works in bands 1-3	The higher the number of works in the smallest bands, the lower the economies of scale and the higher the cost; used in conjunction with the proportion variable to pick up economies of density but could be correlated with load
	Proportion of tertiary activated sludge treatment	Load subject to tertiary activated sludge treatment/Total load	As this is considered the most expensive type of treatment from the ones reported, coefficient sign is expected to be positive. Interpreted against secondary activated sludge treatment proportion
	Proportion of tertiary biological treatment	Load subject to tertiary biological treatment/Total load	Biological treatment is considered less expensive than activated sludge and as interpreted with respect to secondary activated sludge treatment, negative coefficient expected
	Tight consent 3 dummy	Dummy variable for tight consent on suspended solids (SS), five day Biological Oxygen	As tight consent requires companies to meet certain discharge quality, which will

[10] Biological Oxygen Demand.

Type	Variable	Definition/Transformation	Rationale
		Demand (BOD_5) and ammonia[11]; 1 if tight consent exists on both SS and BOD_5 or ammonia	lead to higher opex
	Tight consent phosphorus dummy	Dummy variable for tight consent on phosphates; 1 if tight consent exists[12]	Same logic as consent variable above
	Activated sludge dummy	Dummy variable for the use of activated sludge in large works, 1 if yes, 0 if no	Activated sludge is considered to be the most expensive type of treatment, thus leading to higher cost

3.2.3. Models

There are two issues to consider when thinking about the specific modelling that we undertook:

- First, how was the data combined and over what time period?
- Second, what actual techniques for estimation were employed?

Each is discussed below.

Basic approaches

When thinking about the data questions there were two key ones which would affect the modelling:

- Whether to use pooled and panel datasets
- Whether both base and enhancement capex should be smoothed

A pooled dataset is one where each observation, no matter when it is taken, is treated as being at the same time and for a separate firm. So, effectively there is no time effect, although the models we considered sought to address this by including a time dummy variable which allowed a time trend coefficient to be estimated. A panel dataset takes into account the fact that observations are at different times for a recurring set of companies and that the timing of the observation does matter. The need to allow for time effects (via specific explanatory variables) limits the number of effective observations and so the number of explanatory variables. We considered both in our modelling.

The second data issue arose with respect to whether raw annual capex numbers should be considered (total expenditure). Since capex can be quite variable, especially over a price control period, using a smoothed capex figure – effectively a rolling average – may provide a better measure of expenditure than the actual raw annual figure. This is especially important if the

[11] Thresholds for the determination of consent: 30 mg/l for suspended solids, 20 mg/l for BOD, 5 mg/l for ammonia; if the level of each of these items is below the threshold, tight consent is equal to 1 as lower concentrations are harder to purify.

[12] Threshold: 2 mg/l for works with population equivalent (p.e.)lower than 100,000, 1 for works with p.e. equal or greater than 100,000.

impact on quality etc is expected to occur over time rather than instantaneously in the year in which the capex occurs.[13]

For water activities we tended to use a five year average to smooth capex since even losing four years of observations (to allow for the initial smoothed capex value) still left us with 90 observations. However, we did test the use of raw annual values as well. For sewerage we decided to only use raw values. This was a pragmatic solution to the more limited number of observations available for this activity. By using all nine years of data we were able to also have 90 observations for the sewerage modelling.

Modelling techniques

Within the question of modelling techniques there were two issues that needed to be considered:

- Whether average or frontier efficiency should be considered
- What form of model (raw numbers, log-linear or trans-log) should be considered?

Ofwat has traditionally focused its econometric work on frontier analysis for opex while using average efficiency with the cost base at PR09 to determine capex efficiency, although prior to PR09 its used frontier based approaches for capex as well. In our analysis we have considered both average efficiency and frontier approaches – although the frontier approaches have been econometric ones rather than data envelopment analysis (DEA) type approaches.

We do not think it matters which form of approach is employed, provided that the implications for the other aspects of the regulatory regime are clear. For example, the design of the menu would need to be adjusted if Ofwat used a frontier based approach. Conceptually this is straight-forward to implement (and reflects the approach that Ofgem has applied in its menu design) but it may make the possible changes to the PR14 regime that Ofwat is considering less transparent or more difficult for stakeholders to understand. Consequently we believe that only once the final design of the menu is clearer should any final decision for an average or frontier based approach be made.[14] However, our working assumption for the remainder of this report is that an average efficiency approach will be applied. In part this reflects a presumed desire for consistency within Ofwat – the proposed retail control for households is linked to average costs. Further, our expectation is that the baseline will be based on average efficiency and will be consistent with the previous menu approach adopted by Ofwat.

With respect to the form of model, the choice is about how good a fit the form provides. There is evidence in the sector to suggest that both economies of scale and diseconomies of scale exist (see Stone and Webster 2004 for an analysis of the economies or diseconomies of scale in the water industry). Consequently a trans-log model form is best placed to allow for the changing nature of the economies of scale. Trans-log models are, however, less transparent than other model forms. We also considered log-linear models since they are simpler to replicate but suffer from the imposition of a single form of economy of scale being assumed across the industry – all

[13] One possible area for future consideration would be the inclusion of lagged variables to capture the types of delay or phased effect that can be expected.

[14] This issue is addressed in more detail in CEPA's July 2012 report for Ofwat on incentives and menus, published as part of the August 2012 consultation on wholesale incentives.

companies are assumed to face one of increasing, constant or decreasing returns to scale rather than some belonging to each category which trans-log models allow.

We return to these issues in the concluding section of this report when we consider how our models can be refined and preferred models selected for proposed use at PR14.

4. MODEL OUTPUTS

As explained in the Introduction, our focus is on reporting the results of the modelling at a high level rather than through the detail of the specific models. Water and sewerage are considered separately below.

4.1. Water

Table 4.1 summarises our findings with respect to different aspects of water.

Table 4.1: Summary of findings for water

Value chain element	Expenditure level	Approaches	Comment
Wholesale	Totex	Average efficiency and frontier approaches both provide results in line with expectations Both pooled and panel provide results in line with expectations Log-linear models work for most companies but results may not always be in line with expectations Trans-log models seem more robust than those without varying returns to scale.	A relatively small number of explanatory variables in each of the three categories are statistically significant The refined and broad trans-log models both give expected results which are robust in terms of efficiency levels and order. The log-linear models are not as robust
Wholesale	Base opex and capex	Average efficiency and frontier approaches both provide results in line with expectations Both pooled and panel models provide results in line with expectations Log-linear models work for most companies but results may not always be in line with expectations Trans-log models seem more robust	These models are similarly robust
Wholesale	Enhancement	Raw capex results are more robust than smoothed ones Significance of results is limited and *a priori* expectations are not always met	This is clearly an area of weakness if a more disaggregated modelling approach was to be used as a cross-check on integrated totex modelling results The models can be used but with less confidence than the wholesale and base ones

Given these results it is possible to consider the cost assessment approaches against the criteria established in section 2. Table 4.2 sets out our views on the assessment.

Table 4.2: Assessment of water options against the criteria

Criteria	Assessment
Targeted price control	This links across to the level at which the control is being set. Totex cost assessment seems robust at the integrated wholesale and base levels but less

Criteria	Assessment
	robust at the disaggregated enhancement level
Effective incentives	Wholesale totex cost assessment should limit the opportunities for gaming and so ensure effective incentives The fact that enhancement cost assessment is less robust may introduce opportunities for gaming if a separate base and enhancement approach is adopted as the main approach. This would lead to less effective incentives
Statutory duties and impact	This depends significantly on the approach adopted to use the cost assessment. If a single model were chosen as a preferred approach there is clearly a risk that the proposed revenue allowance could be insufficient for the outcomes required from the company – or alternatively that consumers would be paying too much for these outcomes. The degree of risk will depend on the design of the remainder of the regime – both with respect to risk triggers (for the corridor use of cost assessment) and the menu (for the baseline estimation use of cost assessment)
Implementation	The variables used and the data should be easily available. The need to use the trans-log approach to achieve a comprehensive robust result makes replication difficult without the use of complex econometric packages, ie it cannot be estimated in excel. Further, interpreting the results is less straight-forward owing to the need to interpret the results with respect to economies of scale. Finally, trans-log model coefficient estimates can be volatile when new data, such as an additional year, is introduced. While this may not invalidate the results, it does mean care is needed over the interpretation of point estimates rather than confidence intervals when forecasting with the models

While the assessment is not universally positive, it is also important to bear in mind that the results represent improvements over previous Ofwat approaches in some respects, and what the alternative is likely to be. This issue is considered further at the end of this section.

4.2. Sewerage

Our assessment of sewerage is more mixed than for water. We initially took a similar aggregated approach to sewerage, modelling all expenditure types for the entire wholesale value chain together, which did not yield sufficiently robust results. We briefly present those at the end of this section. These findings were consistent with previous experience of dealing with the more limited data set available for sewerage activities.

We were subsequently commissioned by Ofwat to undertake further work for sewerage to determine whether alternative approaches to cost assessment are more appropriate in this context. As discussed in Section 3, this work focused on:

- further investigation of possible totex models;
- a separation of sewerage activities into sewerage network and sewage treatment (as noted earlier in the report, it was agreed to exclude sludge treatment from this analysis); and
- investigation of possible opex only models (more in line with the models used by Ofwat at the last determination, PR09).

Using further data, including sub company data for sewerage areas and large sewerage treatment works (STWs), provided by Ofwat from the June Returns, the results of our findings for this

additional work are summarised in table 4.4. We have split the sewerage works sub-company data into seven bands based on the relative size of the STWs. The first five bands are the traditional Ofwat size bands as defined in the glossary. We disaggregated the traditional band 6 into large works and mega large works (a new band 7), as the majority of some companies' load is focused in a small number of works. Band 7 works have a population equivalent greater than 342,000.

Table 4.4: Summary of findings for sewerage

Value chain element	Expenditure level	Approaches	Comment
Sewerage network	Base expenditure	Average efficiency and frontier approaches both provide results in line with expectations Both pooled and panel models provide results in line with expectations	While the results of this analysis are robust they are not as robust as those found in the wholesale water work Fact that only base totex has been measured means that some concern about the enhancement boundary will continue to exist
Sewage treatment	Opex	Average efficiency and frontier approaches both provide results in line with expectations Company level analysis not robust for size bands 1-7 but pooled models robust for bands 1-5 Modelling of individual bands provided robust estimates for bands 4-7. Pooled models for these bands worked while panel models for bands 6 and 7 also potentially robust Individual large treatment works (bands 6 and 7) data provided robust pooled models Unit cost approaches employing trend lines give robust results for bands 1-3	While data constraints in relation to capex meant totex sewage treatment models were not possible to investigate, opex only models have provided robust results. It should be noted that for each band of treatment works there are at least two viable methods of cost assessment Clearly the boundary issues between opex and capex continue to raise potential issues, but these are issues that existed at previous determinations
All wholesale	Enhancement	No viable multiple regression approaches that can be used in isolation found	This will force a reliance on alternative approaches such as unit costs or opex econometric models. The implications of this are addressed later in this section

Given these sewerage results it is possible to consider the cost assessment approaches against the criteria established in section 2. Table 4.5 sets out our views on the assessment.

Table 4.5 Assessment of sewerage options against the criteria

Criteria	Assessment
Targeted price control	This links across to the level at which the control is being set. The fact that the most robust results were found for the network aspects means that potentially the cost assessment could be undertaken at a more granular level than the control
Effective incentives	As noted above, it is likely that the cost assessment needs to be undertaken at a level below that of the control. This will clearly introduce the risk of gaming since the network cost assessment is likely to focus on factors that are not part of the enhancement assessment. The effectiveness of the incentives could be affected by this without appropriate mitigations, but this is no worse than the situation that was faced at the last determination

Criteria	Assessment
Statutory duties and impact	This depends significantly on the approach adopted to use the cost assessment. If a single model were chosen as a preferred approach there is clearly a risk of setting an inappropriate allowed revenue level. The degree of risk will depend on the design of the remainder of the regime – both with respect to risk triggers (for the corridor use of cost assessment) and the menu (for the baseline estimation use of cost assessment). With the menu design, the risk is linked to the impact on company returns and their ability to finance investment that would arise from having to select a part of the menu on the right of the break-even point where returns are lower than the expected cost of capital. The company would be choosing this part of the menu because it believes that it needs a higher level of totex than the models were forecasting
Implementation	The reliance on a range of alternative approaches, some of which are at a quite granular level, means that: • risks over boundaries and the possibility of gaming will exist, and will need appropriate mitigations; and • data requirements will be more significant, although at the next review, PR14, these are likely to be no greater in these areas than those that existed at PR09 and probably considerably less onerous for most of the affected WaSCs. With respect to the latter point, for totex models to be considered at the sewage treatment level greater granularity of capex data will be required. This could allow alternative models to be considered for later price reviews, such as PR19. This needs to be addressed as part of the wider PR19 information provision question.

As mentioned above, we analysed sewerage at the aggregate level as well but we do not consider those models to be sufficiently robust for direct use in cost assessment. At that stage of the work, we also explored network only models, which indicated that a disaggregated approach would be more viable. Table 4.6 summarises our findings with respect to these different aspects of the sewerage value chain.

Table 4.6: Summary of findings for failed sewerage models

Value chain element	Expenditure level	Approaches	Comment
Wholesale	Totex and base expenditure	Limited robustness of the models across all approaches	Sludge and sewage treatment differences between companies makes establishing any robust model difficult
Wholesale	Enhancement	Limited data and no robust models found	This is clearly an area where alternative methods need consideration

We also tried a combined wholesale model comprising both water and sewerage. The results from this were better than expected and relatively robust. In principle this may provide a route by which a sewerage estimate could be found – using the more robust wholesale water estimate alongside. While we do not believe this combined service totex approach is sufficiently robust to form the basis of a stand-alone cost assessment we do believe it has a potential role as a further cross-check/robustness check for other cost assessment approaches.

4.3. Summary

In our view, the following conclusions about the viability of totex or other cost assessment approaches arise from our analysis. Note that in this section we discuss viability of econometric modelling for cost assessment at different levels of disaggregation. We have not attempted to identify specific models which should be preferred for cost assessment. This is an element of further work that has been commissioned by Ofwat.

4.3.1. Water

With respect to the viability of totex cost assessment, our findings are that totex assessment at water wholesale level is viable. Our assessment of viability has been based on at least one of our models meeting each of the following: (i) cost drivers which are statistically significant; (ii) the form and 'direction' of coefficients meets expectations (i.e. realistic cost function represented); and (iii) efficiency estimates that seem plausible given Ofwat's previous modelling and determinations. In addition, we note that:

- no single model is preferred;
- unless trans-log models are considered some companies cannot be accommodated easily into a model; and
- at the base opex and capital maintenance level similar results in terms of explanatory variables and values are found to wholesale totex.

In contrast, we found that models which only include enhancement capex are less robust. Therefore, we consider that any standalone assessment of enhancement costs would be likely to require support from other forms of assessment such as unit cost model(s). These results are summarised in Table 4.7 below.

Table 4.7: Viable cost assessment approaches for Water

	Opex	**Base capex**	**Enhancement capex**
Water wholesale (totex model)	Econometric modelling		
Water wholesale (by expenditure type)	Econometric modelling		Non-econometric cost assessment

These models form the basis of our water sector considerations in this report.

4.3.2. Sewerage

We identified models based on the following criteria:

(i) cost drivers which are statistically significant;

(ii) the form and 'direction' of coefficients meets expectations (i.e. realistic production function represented); and

(iii) efficiency estimates seem plausible given Ofwat's previous modelling and determinations.

Given this, our findings with respect to the viability of totex cost assessment are:

- at the overall sewerage wholesale level, totex cost assessment is not a viable approach Ofwat could use in isolation;
- at the disaggregated sewerage network level, totex cost assessment for base costs (opex and capital maintenance) is viable; and
- currently no single model is preferred.

These results are summarised in Table 4.8 below.

Table 4.8: Viable cost assessment approaches for Sewerage

	Opex	**Base capex**	**Enhancement capex**
Sewerage network	Econometric modelling		Non-econometric cost assessment
Sewage treatment	Econometric modelling/ unit costs	Cost base/ unit costs	

Further analysis of alternative cost assessment approaches has led to the following conclusions for sewage treatment:

- viable opex models exist for size bands 1-5, the smaller treatment works, to be considered at a company level;
- an improved approach to opex unit costs is viable for bands 1-3;
- bands 4-7, the medium and large treatment works, can be viably modelled for opex at a band level;
- bands 6 and 7, the largest treatment works, can be viably modelled for opex at a treatment works level;
- where econometric models are being employed, pooled rather than panel models proved to be the most viable; and
- no single preferred models have arisen.

These results are summarised in Table 4.9 below.

Table 4.9: Viable cost assessment approaches to Sewage Treatment opex

Opex	**Econometric model**			**Unit costs**
	Company level	**Company (by band)**	**Sub-company (by works)**	**Company (by band)**
Band 1	✓	×	×	✓
Band 2		×	×	✓
Band 3		×	×	✓
Band 4		✓	×	×
Band 5		✓	×	×
Band 6	×	✓	✓	×

Opex	Econometric model			Unit costs
	Company level	Company (by band)	Sub-company (by works)	Company (by band)
Band 7		✓	✓	✗

Further complications arise with sewerage, in part because of the modelling results, which need consideration. Specifically:

- enhancement capex is an issue throughout sewerage and will need to be addressed through an alternative approach. Ofwat has already started developing models to cover these areas and the initial testing suggests that they should be viable. There are data collection issues, relating both to the quantity and timing of data,that will need to be considered. Other implications linked to this are discussed below;
- capital maintenance for treatment works is being estimated in a different way by Ofwat which again has data issues and implications for the way cost assessment estimates are used;
- sludge treatment and disposal has been excluded from this analysis and requires addressing by Ofwat, which is currently underway, as it will need to be incorporated into the cost assessments that Ofwat undertakes at the next price determination; and
- some new activities, specifically private sewers, have not been considered as none of the available data used to estimate the cost assessment models includes this new activity. The implications of this are considered below.

Note that this report does not make specific recommendations about how these issues can be resolved as this is an area that Ofwat is currently developing its analysis in based on previous Ofwat models, including unit cost approaches.

4.3.3. Implications

What is clear from the analysis we have undertaken is that integrated totex cost assessment as a single exercise for the entire value chain and all expenditure types is only fully viable in water, while it is viable in part in wastewater, with issues and concerns around some areas. Specifically:

- as no single preferred model has arisen consideration of a suite of models is appropriate, this raises questions about how those models are combined but can also be seen as a strength as the range of models allows for triangulation and robustness testing;
- enhancement remains an issue for sewerage and an alternative approach, potentially building on previous methods used by Ofwat, is needed and currently being developed by Ofwat. For water it would be useful to have a robustness test using such an approach, to supplement the more advanced econometric cost modelling;
- aspects of sewerage cost, such as sludge and private sewers, have not been considered, although how they are addressed by Ofwat will remain important; and
- boundary issues may arise given the mix of approaches being used.

Each of these is considered briefly in turn.

Triangulation

When using multiple totex estimates for cost assessment it is useful to think of the concept of triangulation. This is an approach where the results from more than one model are used to derive some form of average estimate. Whether it is just a simple average or one weighted by the degree of robustness of the models depends on the ability to differentiate that degree of robustness. Given that no single preferred totex model has arisen, we think triangulation is likely to be important to further improve confidence in the modelling results.

Alternative approaches for modelling capex

Econometric modelling of enhancement capex has unsurprisingly proven to be more difficult than base costs. This is especially true for sewerage where the available data are very limited and consequently an alternative approach is required. For water, while our totex models provide viable estimates of capex, we believe this should still be supplemented by a further approach as:

- this will provide a further robustness test for the totex estimates; and
- new areas of enhancement capex are unlikely to be captured through the econometric modelling.

It seems that the best option available at the moment is to develop an alternative methodology for the estimation of capex which could build on Ofwat's original cost base and programme costs. As mentioned above, Ofwat has started addressing this issue.

Ofwat's PR09 cost base was a unit cost comparison across about 80 different standardised investment activities which was used to develop an efficiency challenge to the costs submitted by companies in the their business plans.[15]

The existing approach, applied at PR09, was to establish an efficiency challenge based on the relative efficiency of the company that was then applied to the Ofwat baseline. This approach could continue to be used but does raise issues such as timing of data flows from the companies to Ofwat. Specifically Ofwat would not be able to propose an independent estimate of the full baseline for a company until after it had sufficiently reviewed the relevant business plan information, since some of the capex baseline would depend on the business plan outcomes and supporting capex proposed by the company. As per the timetable in Figure 8 in the main consultation document, this should be in early 2014.

Using information contained in the business plan would be a significant reduction in the "truth telling" aspect of the menu, but since water totex, sewerage network base totex, and sewerage treatment opex (at least) would all be independently estimated it does not in our opinion invalidate the use of a menu approach. Robustness checks on the information provided in the business plan – provided both by considering historic information and possible use of

[15] Whether a future version needs to be as comprehensive is something that Ofwat will need to consider (given the decreasing importance of enhancement relative to base capex), but it is clearly important to develop a revised version of this.

independent engineering review (an approach used by both Ofwat and Ofgem in the past), can also supplement the "independence" of the information used by the regulator.

Further, depending on the way that the projects are included in the cost base and business plan, it could be possible to make this more like a unit cost assessment (like the size band 1-5 sewage treatment works was at PR09) which then allows a scaling up of the unit costs to determine a capex requirement (rather than an efficiency challenge). One approach to this could be by seeking unit costs for incremental changes in output or driver levels. Adjustments for regional variations would need to be made, but this should be possible given the information that has been considered for the totex modelling.

While this would still potentially cause a degree of blunting of the independence of the estimate, it would be a solution more in line with the overall approach to cost assessment being proposed in this report. Ofwat is undertaking separate work on capital cost estimation and how this develops will clearly impact on the options that they have for PR14.

Excluded items

Sludge and private sewers have not been included in our examination of the viability of multivariate econometric approaches to cost assessment modelling. The former because of the problems identified in earlier work by Ofwat, associated with the various ways that sludge is treated and disposed of, while for the latter there were no cost data in the existing dataset owing to the transfer of private sewers to the water companies only having taken place within 2010-11.

As discussed in the previous work for Ofwat on Incentives and Menus (published in August 2012 by Ofwat as part of its consultation and addressed in section 5.2.5 of our report) there are several ways in which non-controllable or difficult to estimate expenditure could be handled. These are out of the scope of this report and it will be for Ofwat to decide the right treatment for each type of cost.

Boundary issues

What we mean by "boundary issues" is the problem raised by applying different cost assessment approaches to different elements of the baseline calculation. The ultimate version of this, which Ofwat has been seeking to address through this totex work, is the distinction between opex and capex. Boundaries provide the opportunity for gaming. The sort of gaming that can occur is from exploiting differences in the cost assessment approach and the accuracy of the estimate that was derived so that costs are allocated to the area where the most beneficial cost assessment will occur.

Whether a company can exploit a cost assessment boundary issue depends on the:

- cost assessment approaches applied;
- Ofwat's ability to police the boundary through reporting, auditing etc; and
- ability of the company to predict the baseline.

With the latter point, a combination of triangulation and other tools, incorporating other company data into the cost assessment, should significantly restrict the opportunity for such gaming.

While the viability of totex modelling for water means that this problem should not arise, sewerage is different. Clearly there will be boundary issues in sewerage. It should be remembered that:

- these will be no worse than the ones that existed at the last price determination (PR09);
- enhancement is becoming a relatively less important cost element, in part because of the growth of the existing asset base and in part because the significant investment over the last couple of decades to address issues of quality has decreased as a proportion of total capex;
- the use of base expenditure (totex) for sewerage networks means that there are fewer boundaries than previously existed for these costs; and
- Ofwat's ability to use the price control process to minimise such opportunities.

The last point incorporates several elements, including:

- as noted above, an ability to use the multiple models to determine triangulation approaches that make it difficult for companies to predict the cost assessment outcome;
- the risk based review of business plans which will allow various triggers to be employed that can prompt a more detailed review of the plan based in part on understanding of the relevant costs derived from other companies' information; and
- multiple approaches for the direct estimation of opex which means that Ofwat will be in a stronger position to understand the direction of the gaming opportunities.

5. BUILDING A CORRIDOR

5.1. Background

Having determined that totex cost assessment is viable, albeit through the use of disaggregated approaches in sewerage, we can now consider how it could be employed in the first of the two cost assessment uses – that of building a cost corridor. The totex cost assessment from the more advanced econometric models would be used in conjunction with Ofwat's other modelling – including its unit cost modelling to build a picture of the overall expediture concerned.

As described in the Introduction, Ofwat's more risk based approach to assessing business plans will require them to have measures or tests that can be employed to determine whether a more detailed assessment is triggered. Clearly the proposed revenue requirement, based on the totex requirement of the company, will be an important element of any such assessment.

While the precise way in which the cost assessment test will be designed will depend on the other tests put in place, this section provides some examples of how such a test could be calibrated and illustrates them through the use of some of the cost assessment models developed. We focus on the wholesale water models as an example of the options. We note that these approaches are examples of how Ofwat could address these issues and does not exhaust the list of indicators that Ofwat might want to use for cost assessment.

The design of other triggers is outside the remit of this paper, but there are clear links. For example, if the robustness of the cost assessment models was deemed to be poor, or they did not collectively cover all the relevant costs, it could be appropriate for Ofwat to introduce further triggers that help identify gaming opportunities in relation to other costs, or the risk of misclassifying a company.

Finally, we concentrate on establishing a single cost corridor in this section – whether it is for water or sewerage. In principle, especially in sewerage where cost assessment may happen at a more granular level, a cost corridor for each segment could be developed. This would clearly lose any ability for robustness to arise from a portfolio effect of adding the cost assessment derived estimates together into a single cost corridor. However, it is an option that Ofwat might choose to consider depending on the relative robustness of the estimates concerned.

5.2. Approaches

When building a cost assessment corridor there are questions relating to:

- the symmetry of the corridor;
- the way in which it is formed; and
- the relative range of the corridor.

Each is discussed in turn below.

5.2.1. Symmetry

Is what we are interested in a corridor or a ceiling? If it is a corridor then we are effectively thinking about something where both the floor and ceiling would trigger further review. With a ceiling we are only worried about whether the value is too high. In part the answer to this lies with whether Ofwat would be concerned that a company which significantly under-forecast its costs (relative to what Ofwat believes to be required) could pose a risk to its statutory objectives (e.g. the financeability of efficient companies and the bills faced by future consumers) in the future.

The choice could also depend in part on the form of cost modelling that has been undertaken. Relative or average efficiency modelling could lend itself to symmetric corridors while ceilings may be more suited to frontier based modelling.

We consider both symmetric corridors and ceilings in our illustrations below.

5.2.2. Formation

There are many ways in which a cost corridor could be formed, but here we consider two as examples:

- a chosen percentage range centred on a triangulated average arising from the suite of cost assessment models;[16] and
- an upper and lower value determined by the highest and lowest totex estimates arising from the suite of models – effectively a maximum and minimum based corridor.

Of course, it is also possible to scale the maximum and minimum (max-min) values if a greater (or smaller) range for the corridor was required.

5.2.3. Relative range

How wide a range should there be for the corridor? This will depend on:

- the degree of confidence in the models;
- the other triggers or tests of the business plan being employed by Ofwat; and
- the design of the menu being used to incentivise efficient delivery by the company.

There is also a potential link between the degree of confidence in the model and the design of other possible triggers. Consequently some of the points above are not mutually exclusive.

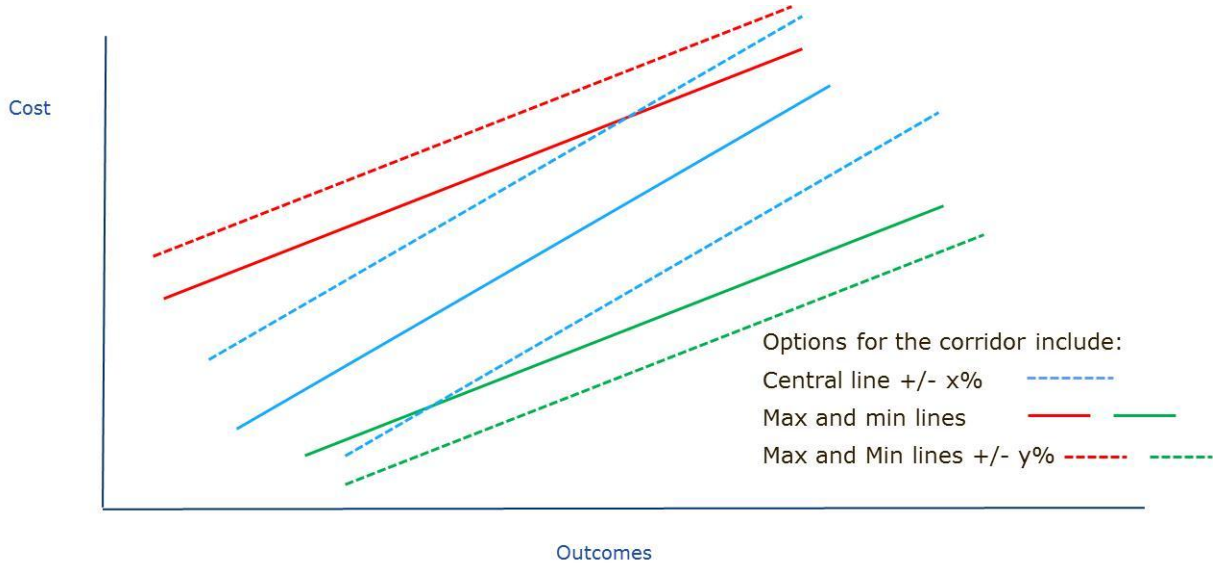
Clearly too great a range will raise the possibility that companies are proposing cost forecasts that are actually inefficient while too small a range may create the risk that actually efficient proposals are unnecessarily reviewed.

5.2.4. Examples

[16] It is possible that this is not the estimate of the baseline as, in principle: Ofwat might choose to use different models when setting the cost corridor and the baseline (the latter might be based on a subset of the former).

Figure 5.1 provides an illustration of what different cost corridors could look like. Both a chosen range around a central, triangulated, estimate as well as max-min approaches (raw and scaled) are shown. This is based around an average efficiency approach.

Figure 5.1: Examples of possible cost corridors



Obviously the illustration is a simplification but it conveys the idea of how corridors could be formed.

5.3. Examples of possible corridors

Using four of the water totex models we identified, we tested how many of the 18 water companies would have had their water cost forecast within the corridor if the last year of data from our dataset was used as the forecast. The results of this, using various mixtures of symmetry, ranges and approach are shown in table 5.1.

Table 5.1: Water totex average efficiency trigger ranges

Measure	**No of companies within range**
Max to min estimated value range	7
Within y% of:	
Average of all totex estimates	5
Average of all non-translog totex estimates	7
Average of all translog totex estimates	7
Within x% of:	
Average of all totex estimates	10
Average of all non-translog totex estimates	11
Average of all translog totex estimates	11
Below the maximum estimate value	13
Below 1+y% of:	
Average of all totex estimates	12
Average of all non-translog totex estimates	13
Average of all translog totex estimates	13

Below 1+x% of:	
Average of all totex estimates	13
Average of all non-translog totex estimates	16
Average of all translog totex estimates	14

Overall, it would seem that key elements are:

- the impact of choosing a ceiling or a symmetric approach is significant; and
- unsurprisingly, how great a range is allowed.

There is a clear interaction between these two elements. The impact of a wider range, using these illustrative models, is much less significant when a ceiling rather than corridor approach is employed.

For completeness we also considered what happened with our frontier models. Table 5.2 provides the illustrative results for all three of the cost levels we estimated, not just totex.

Table 5.2: Water frontier cost estimate ranges

Number of companies with actual lower than:	**Totex**	**Base**	**Enhancement**
Maximum frontier estimate	1	2	3
Minimum frontier estimate	0	0	
1+y% of maximum	4	4	3
1+y% of minimum	1	0	
1+x% of maximum	9	5	4
1+x% of minimum	3	1	

Note: as only one enhancement model has been chosen the max-min values are the same.

Smaller numbers of companies are within the corridor, especially when compared to the ceiling approach in table 5.1 above. This suggests that if frontier based approaches are finally chosen as the preferred way forward a greater distance to the frontier has to be allowed for the acceptable range.

5.4. Summary and assessment of the approaches

A formal assessment of the different approaches to forming the cost corridor is not appropriate at this time as it is dependent on a range of factors that have not yet been determined and requires final versions of cost assessment models to be chosen. However, we believe that the illustrations in this section show that this approach is viable if Ofwat were to decide that it wishes to use cost corridors as indicators. If it does decide to use cost corridors, the key questions for Ofwat will relate to:

- whether cost forecasts that are too low also require further review;
- the size of the acceptable range; and
- whether there are implications for the menu design.

While providing an answer to the first question would seem to be appropriate, it may be that Ofwat does not want to provide an answer to the second. If Ofwat is too explicit it runs the risk that companies will adjust their business plans to exploit the implied opportunity to maximise the cost forecast/revenue requirement without triggering a more detailed review. As such, Ofwat should ensure that the ability to predict the baseline and the associated corridor levels is limited and/or ensure that other business plan tests provide sufficient discretionary power if bunching of forecasts is seen. If there is sufficient confidence that the baseline is more difficult to predict then Ofwat should provide greater certainty to companies about the proposed width of the corridor.

Related to this second point is the question of how Ofwat should develop a cost corridor if it is placing greatest reliance on only one model. While it may be possible to determine a single preferred model we:

- expect other models to continue to be considered as part of a robustness check;[17] and
- absent further updated data, are uncertain whether it will be possible to choose a single preferred model.

If it is possible to find a single model that can be employed as the main or only assessment tool, with sufficient confidence, then how might a cost corridor be determined? Clearly there are different approaches which could include using:

- the confidence interval created by the model estimation for the explanatory variables and use that to derive a corridor (here it would be possible to use either a 50% or 95% confidence interval type approach);[18]
- simple measures around the average value generated for a company (particularly if confidence intervals are not available around the average), as illustrated earlier in the section, say applying a y% or x% scalar to the average; or
- more complex measures around the average value, possibly generating wider corridors at the extremes (this could be relevant if the model is unable to handle the smallest and largest companies but works well for "average" companies).[19]

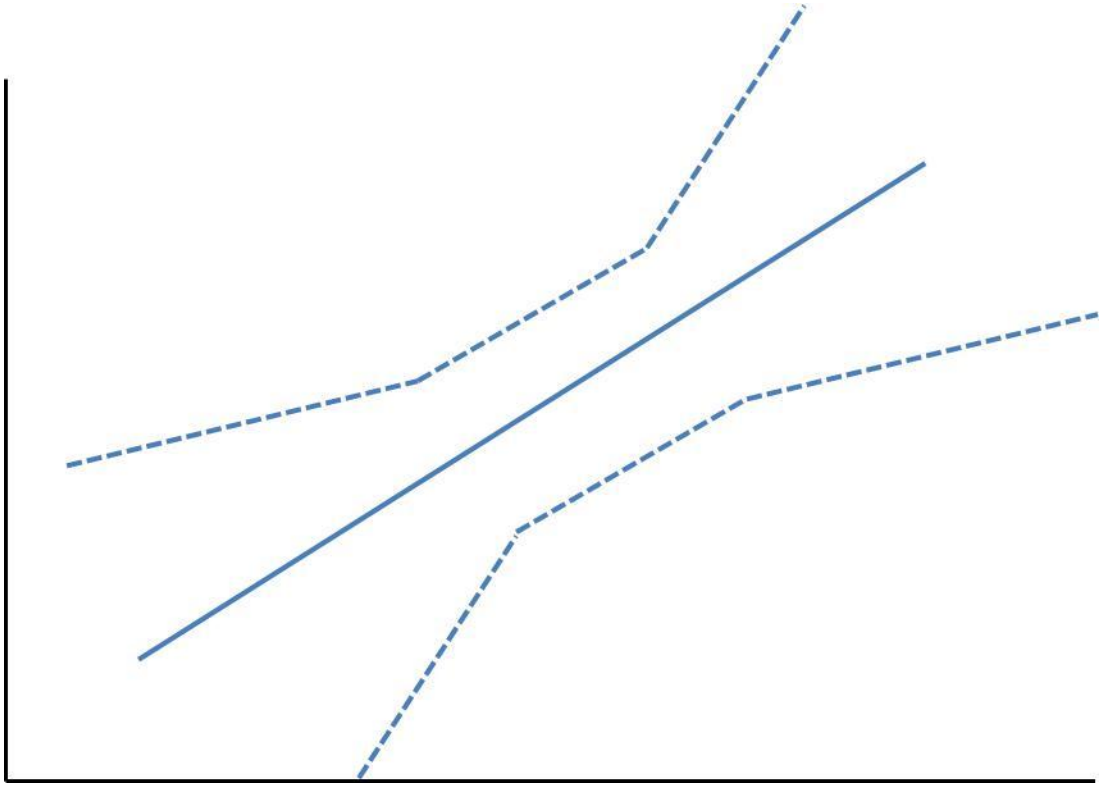
Figure 5.2 provides an example of what the latter type of corridor could look like.

[17] There is a concern that running multiple models will require additional data. This is possible but is unlikely to be a material issue. What is more likely is that different forms of equation (log-linear versus trans-log etc) or full and simplified models (where the latter just uses a subset of variables from the former) continue to be considered. We are, however, aware of the data cost concerns and that should be an element of any refinement/selection criteria used for future work. This is discussed further in section 7.

[18] There is clearly a risk of using all the lowest estimates and all the highest estimates to derive a corridor. The probability that all the highest or lowest values would coincide is likely to be very small meaning that the range derived from this approach is unduly large. Further consideration of this is needed.

[19] This approach could be generalised into an approach based around a different width of corridor depending on some scale factor etc.

Figure 5.2: corridor capturing the greater uncertainty around non-average companies



A final issue to consider is whether there are menu design implications arising from the corridor design. If a wide corridor is used, owing to uncertainty over whether the independent assessment is a reliable estimate of efficient expenditures, which then means a significant number of companies might have their own business plan forecast accepted as the baseline - what does this mean for menu design? The answer is not simple. For example, one interpretation would be the need to have:

- low incentives around the baseline as a way of ensuring that any out-performance is worth little to the business (to limit the incentive for a company to seek to place itself at the upper edge of the corridor in preparing its business plan, with associated high potential rewards from outperformance); or
- high incentives around the baseline so that any under-performance is significantly penalised as this could imply that the company is very inefficient (if it has chosen to already place itself at the upper end of the corridor).

Clearly, one solution would be an asymmetric menu that had different rates for over- and under-recovery. This type of approach has been considered in other industries and rejected as being unduly complex. This in part reflected the overall perception that even the standard menu was difficult for stakeholders to understand and that companies may not respond to the incentives being created. While we do not think the concept of asymmetric incentives is too complex there are clearly issues around transparency and implementation that could arise, especially if capex estimates are smeared or smoothed across the years when actual capex would be different. The loss of symmetry could then create perverse incentives for the company or require some form of adjustment to control for the smoothing.

Further consideration of this issue is required. Once the likely relative size of the corridor is identified, it should be possible to consider the associated implications for consistent menu design.

6. ESTIMATING A BASELINE

6.1. Background

The second use for cost assessment is in setting the baseline to be used for the future price determination. Again, as with the cost corridor, we use this to illustrate what is possible with the wholesale water totex models. Further examples are provided in an annex.

6.2. Approaches

Our focus throughout this report has been on the direct estimation of totex and we continue this here. Also, we continue to assume the use of triangulation as no single preferred model exists. As such, we can see two simple approaches to estimating the baseline:

- the triangulated estimated value; and
- the minimum of the triangulated estimated value and the company's forecast value (proxied by the actual value).

With both approaches it is possible to introduce a scalar into the estimated value to allow for the fact that a company need not be efficient. This is similar in concept to the cost corridor range discussed in Section 5 above.

6.3. Examples of possible baseline estimates

Using three wholesale water totex models, we are able to consider two possible alternative ways of setting the baseline. These are using the:

- estimated values from the triangulation process; and
- minimum of the company's BPF and 100% of the estimated values from the triangulation process. [20]

Table 6.1 summarises the results from these calculations.

Table 6.1: Results of different options for setting the baseline

Approach	Value (£m 2010/11 prices)	% difference to actual
BPF	4,130	-
Estimated	4,064	2
Minimum of BPF or 100% of estimated	3,891[21]	6

As can be seen, the choice of approach leads to small but not insignificant reductions compared to the business plan forecast. It should be noted that these are purely illustrative – the values to which the approach is being applied are themselves the results of the PR09 efficiency analysis and consequently not a fair reflection of what might be achieved.

[20] For this example we have used actual values as the business plan forecast.

[21] Calculated as the minimum of the forecast or Business Plan value.

6.4. Summary and assessment of the approaches

This brief illustrative discussion of the exact way in which the cost assessment models could be used to set the baseline should show that the approach is viable. What will matter will be:

- the degree of self-scrutiny applied by the companies through the business planning process; and
- the design of the menu and incentive systems by Ofwat.

Only once the latter, and preferably the former, are known would it be possible to calibrate exactly how the baseline estimation should be undertaken. We understand that Ofwat will decide on these issues following consultation over the practical use of menus in the price review.

A further issue, linked in part to the implementation of the menu and incentive system by Ofwat, which needs to be considered is how baseline setting might differ if Ofwat produces "draft" and "final" baselines between its draft and final determinations. This will depend on the timeline developed by Ofwat.

7. CONCLUSIONS AND SUMMARY

This report has set out the findings from the project we have undertaken for Ofwat. The remit of the project was to test the viability of total cost or totex assessment for PR14 and beyond. We have done that and also:

- checked to see whether there are obvious problems with implementing the two potential uses of cost assessment at PR14; and
- for sewerage, where totex cost assessment of the whole value chain in single models is problematic, considered possible alternative more disaggregated cost assessment approaches.

7.1. Cost assessment

We have assessed the models against the following criteria:

(i) cost drivers which are statistically significant;

(ii) the form and 'direction' of the coefficients that meets expectations (i.e. realistic production function represented); and

(iii) efficiency estimates that seem plausible given Ofwat's previous modelling and determinations.

Based on this assessment the following conclusions about the viability of totex or other cost assessment approaches can be drawn.

7.1.1. Water

With respect to the viability of totex cost assessment, our findings are:

- at the wholesale level totex cost assessment is viable;

- no single model is preferred; and
- unless trans-log models are considered some companies cannot be accommodated easily into the models.

These models form the basis of our water sector considerations in this report. As discussed in the main report, there may be some benefits from testing the robustness of aspects of the enhancement capex estimate through a consideration of the alternative cost assessment approach being considered by Ofwat. However, this would be just a test of robustness and consequently should only be considered if the costs of data gathering and risks to the truthfulness of business plans are limited.

7.1.2. Sewerage

With respect to the viability of totex cost assessment, our findings are:

- at the aggregated level of the whole wastewater service, integrated totex cost assessment is not a viable approach; but
- at the sewerage network level, totex cost assessment for base costs (opex and capital maintenance) is viable; and
- for this purpose, currently no single model is preferred.

Our analysis of alternative cost assessment approaches for other components of wastewater costs has led to the following conclusions for sewage treatment:

- viable opex models exist for size bands 1-5, the smaller treatment works, to be considered at a company level;
- an improved approach to opex unit costs is viable for bands 1-3;
- bands 4-7, the medium and large treatment works, can be viably modelled for opex at a band level;
- bands 6 and 7, the largest treatment works, can be viably modelled for opex at a treatment works level;
- where econometric models are being employed it has tended to be pooled rather than panel models that have proven viable; and
- no single preferred models have arisen.

In relation to other sewerage costs, in the light of the more advanced econometric modelling results, alternative approaches need consideration. Specifically:

- enhancement capex is an issue throughout sewerage and may need to addressed through an alternative approach, possibly some form of unit cost analysis. There are data collection issues, relating both to the quantity and timing of data, that will need to be considered to apply such an alternative, similar to challenges arising in previous reviews for these expenditures. Other implications linked to this are discussed below;

- capital maintenance for treatment works will also have to be estimated in a different way which again has data issues and implications for the way cost assessment estimates are used;
- sludge treatment and disposal has been excluded from our econometric analysis and requires addressing by Ofwat as it will need to be incorporated into the cost assessments that Ofwat undertakes at the next price determination; and
- some new costs, specifically private sewers, have not been considered in our econometric analysis as none of the available data used to estimate the cost assessment models includes this new activity. The implications of this are considered below.

Ofwat has been developing alternative ways to undertake cost assessment for these areas.

7.2. Corridors

There are various ways in which the cost corridor could be set. The examples used in Section 5 of the report demonstrate that relatively simple definitions of a corridor could be used to select (at least on the cost dimension) whether a business plan should be subject to further scrutiny. What will be important for Ofwat is to consider the relationship between the tests of the different dimensions so that the right calibration of cost test can be implemented.

The way that the cost corridor is developed will depend on the number of models that are chosen for final development. Ofwat will also need to consider how much information it provides on its chosen approach to cost corridors given the risk of gaming that could arise. Consequently ensuring some regulatory discretion in this area will be important.

7.3. Baselines

Clearly the need for triangulation means that there are several possible ways in which Ofwat could set the baseline. Again, the relatively simple examples in Section 6 demonstrate that even using the existing data it is possible to see how different options for setting the baseline produce slightly different results. Again, the key decisions will relate to how Ofwat wants to set the baseline and this will depend importantly on the design of the menu.

7.4. Next steps

While this report has shown the viability of totex modelling, there are clearly some issues to be addressed about aspects of this. In part this can be overcome through the use of multiple models, or suites of models. But this is not sufficient for some aspects, especially relating to sewerage, where opex only models are likely to be used at least for sewage treatment. As such, we believe the work on alternative approaches to capex, that Ofwat has been developing will be an important complement to the econometric models discussed here, especially:

- capital maintenance for sewage treatment, sludge and sludge disposal; and
- enhancement capex for all sewerage.

In addition, supporting the water enhancement capex assessment through the use of approaches previously used by Ofwat , including unit cost models, will help ensure the robustness of the results with respect to changing outputs.

Further work is also required to refine the models that have been identified as being viable. This work is likely to focus on assessing the:

- theoretical "correctness" of any variables and their coefficients;
- the theoretical validity of the model form (economies of scale etc);
- relative fit of the models; and
- practical implementation issues around the models including their replicability in the future (linked to requirements for PR19).

As noted in section 5, one of the implementation issues that need to be considered is the cost implications for companies and Ofwat of multiple models. If additional data is required then the incremental costs of having more than one model need to be considered against the incremental benefits of having additional approaches that can provide robustness checks. As noted earlier, we believe there is significant benefit in multiple models and consequently the data collection costs would need to be material for us to recommend dropping models.

This should lead to a small number of preferred models, or even possibly one model if there were very substantial data issues as noted above, for each area of cost assessment, specifically:

- wholesale water (totex);
- sewerage networks (separate base and enhancement capex); and
- sewage treatment (separate opex and capex).

We believe that the work undertaken to date provides a universe of models from which a handful at most of new models might need to be run where either explanatory variables are substituted, dropped or added.

A further two years' information should be available for additional refinement or updating of the models later in 2013 (that is Ofwat can collect the missing data for 2011/12 and a good forecast for 2012/13).

ANNEX 1: OFGEM RIIO COST ASSESSMENT MODELLING

A1. Ofgem's (RIIO-GD1 Initial Proposals) preferred cost assessment approach

As part of its new RIIO (Revenue = Incentives + Innovation + Outputs) regime Ofgem's proposed to use a suite of cost assessment tools assess the regulated companies' efficiencies and business plans. In its original cost assessment proposals, Ofgem indicated that econometric modelling could be conducted at three different levels of cost disaggregation:

- activity level (or bottom-up);
- expenditure type (or middle-up) – capex, replacement expenditure (repex) or opex; and
- totex.[22]

Ofgem also indicated, in its original proposals for undertaking cost assessment, that it could use both historical and forecast data in its models.

In its Initial Proposals for RIIO-GD1, Ofgem decided not to use middle-up assessment in setting the efficiency targets as it gave broadly the same comparative efficiency scores for GDNs as the totex models. Ofgem also concluded that the use of forecast data for the entire duration of RIIO-GD1 (eight years) was not plausible due to the different assumptions used by the GDNs in order to forecast business plans. Ofgem found that the use of historical data gives more reliable estimates and hence is preferable over forecasted data.

There were a number of cost areas which Ofgem excluded from the regression analysis in order to assess the companies on a consistent basis. These costs included street works' costs, TMA (transport management act penalties), smart metering costs, etc. For such costs, Ofgem carried out an assessment on the efficient level of costs based on a technical or engineering assessment. In addition, Ofgem's econometric models excluded costs which Ofgem considered were non-controllable such as network rates and licence fees.

Ofgem's preferred models for the GDNs were: two totex – one based on three years historical data and one based on two years of forecast data – and two activity level models totex – one based on three years historical data and one based on two years of forecast data. The rationale provided by Ofgem for preferring these models is that they gave more or less consistent rankings of GDNs in terms of efficiency performance. The modelling therefore gave a total of four different efficiency scores for each licensee. The simple average of the four efficiency scores was used to get the total % reduction in the GDNs' forecasts for outputs and efficiency.

A1.1. Econometric modelling

Ofgem's econometric model was based on an OLS estimator with a Cobb-Douglas functional form for its econometric model and using panel data fixed time effects. The historical data models used three years of data from 2008-09 to 2010-11 with 2010-11 being the base year. The forecast data was for two years, 2013-14 and 2014-15 with 2013-14 being the base year. The output driver was based on a composite scale variable (CSV) which combines the network scale

[22] Ofgem, Decision on strategy for the next gas distribution price control - RIIO-GD1 Tools for cost assessment, March 2011.

based on modern equivalent asset valuation (MEAV) with workload drivers based on bottom up regressions. The weighting were: 38% on MEAV, 43% on repex workload, 2% each on mains reinforcement and connections workload, 6% on external condition reports, 5% on maintenance MEAV and 4% weighting on the emergency service CSV. Ofgem considered that this reflected a balance of fixed and variable costs. Ofgem made adjustments to the GDN's historical and forecast data for:

- disallowed outputs;
- GDNs forecast workloads;
- to ensure consistency in cost allocation and exclusion of costs addressed through uncertainty mechanism; and
- regional adjustments.[23]

Each GDN's totex efficiency score was calculated as a ratio of its adjusted historical or forecast totex to its modelled totex. The efficient costs were set equal to the upper quartile (UQ) costs, implying that each licensee is required to 'catch-up' to the performance of the upper quartile. The allowances are to be based on the expectation that the GDNs could close 75% of the assessed gap between their forecasts and the upper quartile performance.

Once the efficiency scores were calculated in the base year (2014 for forecast data), the efficient base year scores are then rolled forward to derive the cost allowance for the rest of the years of RIIO-GD1 by adjusting for changes in workload, real input price growth (RPEs), ongoing efficiency, outputs etc as proposed by the GDNs.

The bottom-up analysis included both regression analysis and qualitative and technical assessment of costs. The econometric modelling was carried out for several activities, namely:

- work management;
- emergency service;
- repairs;
- maintenance;
- mains reinforcement;
- connections; and
- Tier 1 repex.

Like the totex benchmarking the efficient base year for historical data in 2010-11 and for forecast data was 2013-14. The efficiency scores were calculated as a ratio of the sum of controllable, normalised and adjusted costs to the sum of the modelled costs for each of the activities. Each licensee's costs were benchmarked against the upper quartile. The efficient base year cost was rolled over in the same way as for totex to calculate the cost allowance for the entire price control period.

[23] Ofgem, separately, set adjustment factors for real price effects (RPEs) and ongoing efficiency, with the former being dealt with via an ex ante allowance.