WSX21 - Annexes - Residential retail

Business plan 2025-2030

Wessex Water
YTL GROUP

FOR YOU. FOR LIFE.

WSX21 - Annexes - Residential retail

CONTENTS

A1 Residential retail cost assessment at PR24: econometric benchmarking models 1

A2 Review of Ofwat's approach to inflation indexation for residential retail price controls 2

A3 Review of Ofwat's use of water companies' cost forecasts in its determination of allowances for residential retail costs 3

A4 Residential retail cost assessment at PR24: projection of cost benchmarks 4

This supporting document is part of Wessex Water's business plan for 2025-2030.

Please see 'WSX00 – Navigation document' for where this document sits within our business plan submission.

More information can be found at

A1 Residential retail cost assessment at PR24: econometric benchmarking models

Residential retail cost assessment at PR24: econometric benchmarking models

Report for Bristol Water
and Wessex Water – 6 April 2023

Reckon LLP
Data analysis | Economic regulation | Competition law

Table of contents

Executive summary ... 3
1: Technical summary ... 5
2: Approach to developing and reviewing models ... 17
3: Review of Ofwat's PR19 residential retail benchmarking ... 23
4: Models of bad debt related costs ... 37
5: Models of other retail costs ... 50
6: Disaggregated versus aggregated models ... 64
Appendix 1: Enhanced metrics and techniques to support model assessment ... 70
Appendix 2: Further information on data used ... 77
Appendix 3: Results from re-running PR19 models ... 79

Executive summary

Background to the report

In the context of Ofwat's PR24 review to set price controls for the period from 1 April 2025 to 31 March 2030, Bristol Water and Wessex Water commissioned Reckon to carry out a project concerning the costs of residential retail activities. The project involved three closely related objectives:

1. To review Ofwat's approach to cost assessment for residential retail activities at PR19, including the set of econometric benchmarking models that it used, and identify ways to make improvements for the purposes of the PR24 review.
2. To develop a suite of econometric benchmarking models for water companies' residential retail activities that might be used for PR24 and to estimate these using the latest available data.
3. To provide analysis and evidence to help Bristol Water and Wessex Water make projections of the efficient costs of their residential retail activities over the 2025-30 period, drawing on the suite of econometric models developed as part of the project.

This report is focused on econometric benchmarking models for water companies' residential retail activities. It provides a review of Ofwat's models and modelling approach from PR19 and sets out alternative models that could be used for PR24.

Ofwat's PR19 models of residential retail costs

For PR19, Ofwat used a suite of econometric models for the purpose of benchmarking water companies' residential retail costs and setting allowances. Its use of econometric modelling was a clear improvement on the approach it had taken at PR14, when it had set allowances for residential retail activities using unit cost metrics, with some separately calculated adjustments, rather than via an econometric benchmarking approach.

The econometric models used at PR19 can be seen to work, in the specific sense of providing a more sophisticated and reliable calculation of retail cost benchmarks than the unit cost metric that had been used at PR14. Unlike the unit cost benchmarks, the PR19 models take account of a number of relevant cost drivers beyond company scale. Nonetheless, the suite of models used at PR19 has a series of significant limitations. We expect that a considerable amount of the differences between companies' actual costs and the benchmarks from the econometric modelling will reflect modelling limitations and noise in the data rather than genuine differences in efficiency between companies.

Exploration and development of alternative models

Drawing on our review of Ofwat's econometric models at PR19, and on our broader expertise, we agreed with Bristol Water and Wessex Water the areas of model specification and ideas for potential improvements that we would focus on when exploring models that might be used at PR24.

We used a range of information, evidence and analyses to assess the relative merits of alternative model specifications and to compare these against the models from PR19. In doing so, we drew in part on the types of established metrics and tests that Ofwat has used in the past and which it told companies to consider for the purpose of submitting their econometric models to Ofwat in January 2023. We also drew on some enhanced metrics which Reckon has developed as a means to tackle some of the limitations of the techniques used by Ofwat and to bring a more reliable and insightful evidence-base for the purposes of developing and selecting benchmarking models.

The main outputs from our model development process are as follows:

- Insight and evidence on econometric modelling approaches that might be used for benchmarking residential retail activities at PR24.
- A suite of econometric models for residential retail activities, estimated using updated data, covering years to 2021/22, which we consider to be an improvement on the set of models used by Ofwat at PR19.

The project was not intended to cover all aspects of model development and analysis that might be used for the PR24 review. The models and modelling approaches arising from it reflect the issues that we agreed with Bristol Water and Wessex Water to attach greatest priority to and reflect the data available for the project. Nonetheless, the outputs should make a significant positive contribution to the cost assessment process at PR24.

Overview of this report

The main body of this report is organised as follows:

- Section 1 presents a technical summary of our econometric modelling, including the suite of econometric models that we developed as part of the project.
- Section 2 provides an overview of the approach we took to developing and reviewing models for the purposes of benchmarking water companies' residential retail costs.
- Section 3 presents our review of Ofwat's models and modelling approach from PR19.
- Section 4 presents key aspects of the rationale and evidence that support the models and modelling approaches for bad debt related costs that feature in our model suite.
- Section 5 presents key aspects of the rationale and evidence that support the models and modelling approaches for other retail costs that feature in our model suite.
- Section 6 presents models of total residential retail costs and provides evidence concerning the relative merits of these compared to the disaggregated models covered in sections 4 and 5.

This report includes three appendices. Appendix 1 explains the enhanced metrics and tools we have drawn on. Appendix 2 provides further information on the data sources and variables used in our modelling. Appendix 3 presents results from re-running the PR19 models using the updated cost assessment data published by Ofwat in November 2022.

1: Technical summary

This section presents a more detailed and technical summary of our econometric benchmarking work. It covers our approach, main findings, and the specifications and results for a suite of models that we developed. In this summary, we take the following topics in turn:

- Our review of Ofwat's PR19 econometric models.
- Prioritisation of the opportunities for modelling improvements.
- Approach to model assessment and review.
- Aggregated versus disaggregated models.
- Models of bad debt related costs.
- Models of other retail costs.

This report is focused on benchmarking models for residential retail cost assessment. It does not consider the retail costs of suppling non-residential customers. Wherever we refer to total retail costs, bad debt related costs, other retail costs, or billed revenue, we are referring to costs or revenues that companies report for their supply to residential customers.

Our review of Ofwat's PR19 econometric models

The use of econometric benchmarking models for price control cost assessment is an ongoing, evolutionary, process through which improvements can be made over time. There is not necessarily a target endpoint of a "robust" set of models, or a correct modelling approach. But it is reasonable for water companies and other stakeholders to expect improvements from one price control to the next.

Our approach in this project has been to take the set of models that Ofwat used in its PR19 final determinations as a starting point and to carry out a detailed review of these models, as a means to understand the limitations and drawbacks of those models and to help identify potential opportunities for improvements.

Turning to some specific points from our review of the PR19 models, we highlight the following:

- **Fit with the data**. We can see limitations in models by looking at R-squared measures and the spread of efficiency ratio across companies (though care is needed in making inferences from these). The models of bad debt cost have a reasonably high R-squared yet still lead to a wide range of efficiency ratios across companies, which suggests cost drivers are not fully accounted for. The models of other retail costs have a very low R-squared indicating that these models explain only a small proportion of the variation in these costs across companies and over time and also show quite a wide range of efficiency ratios.
- **Cost drivers**. Ofwat's retail models provide a reasonably good coverage of cost drivers for retail costs, capturing a number of different factors in at least some of the models in the suite. This is not to say that all plausible cost drivers were taken into account, and we have identified

some potential gaps. Furthermore, we have some concerns about how well the explanatory variables used in the model specifications can capture the underlying cost driver relationships.

- **Level of aggregation of retail modelling**. In contrast to its approach to wholesale cost assessment, Ofwat gave a greater weight to results from its aggregated models (covering total residential retail costs) compared to those from its more disaggregated models (bad debt related costs and other retail costs taken separately). We do not think that there was a good basis for this aspect of Ofwat's approach at PR19. Ofwat's published documents do not seem to provide *valid* evidence on the relative merits of its aggregated models. On the results that it does present, these have some significant disadvantages compared to the disaggregated models.

- **Model dynamics**. Ofwat's PR19 models were static in the sense that the time dimension was entirely absent from their specification. All of the model specifications included a constant term rather than any dynamic elements such as year-specific dummy variables or a time trend. The model specifications imply an assumption that, other than through the effects from changes over time in the cost driver variables (e.g. increases in meter penetration), water company costs do not change (relative to CPIH) over the period 2013/14 to 2018/19. We consider this a shortcoming of these models, especially given other evidence of changes in the levels of costs over this period.

- **Treatment of depreciation and recharges**. We identified two concerns with the way that Ofwat treated the data on depreciation and amortisation in the modelling. First, Ofwat adopted an approach of smoothing depreciation over the data period, by replacing annual depreciation figures with an amount averaged over a five-year period. We did not see a good justification for adjusting companies' reported depreciation data in this way. Second, Ofwat's approach allocates 100% of depreciation and amortisation costs to "other retail costs", when we would expect some to be attributable to debt management activities and so captured, in the modelling, by the models of bad debt related costs. This second issue also applies to recharges: Ofwat's approach allocated 100% of these to the model of "other retail costs" but some of the costs reported as recharges will be costs that are attributable to debt management activities.

We discuss the findings from our review in more detail in section 3 of this report.

Prioritisation of the opportunities for modelling improvements

It is almost always possible to improve upon an existing set of models with sufficient expertise, data and time. We have sought to take a proportionate approach in this project, looking to explore and take opportunities for improvement against the PR19 models without seeking to exhaust all reasonable lines of potential model development or refinement.

Drawing on our review of Ofwat's econometric models at PR19, and on our broader expertise, we agreed with Bristol Water and Wessex Water which areas of model specification, and which ideas for potential improvement, to focus on when exploring alternative models – and modelling approaches – that might be used at PR24. We prioritised the following:

- We explored alternative approaches to specifying the model dynamics, considering alternatives to the static modelling approach used by Ofwat at PR19.

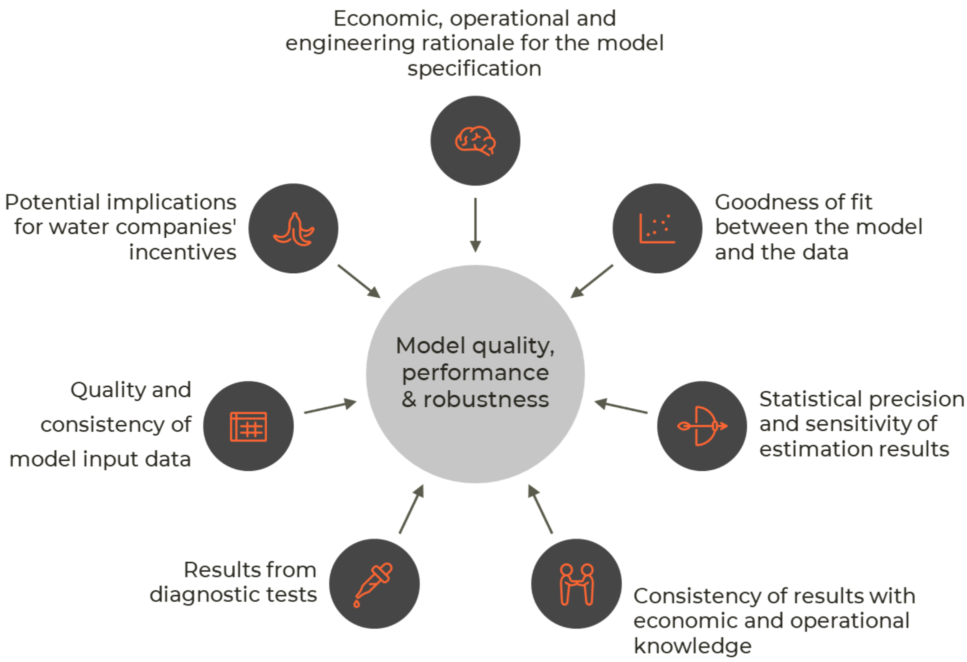
- We considered aggregated models of total residential retail costs and disaggregated models for bad debt related costs and other retail costs taken separately, and we sought to assess the quantitative evidence as to whether one type performs better overall than the other.
- We explored different ways to specify the dependent variable in the econometric models of residential retail costs (e.g. for models of bad debt related costs, we saw merit in considering models where the dependent variable is expressed in terms of the ratio of bad debt costs to billed revenue, in addition to the PR19 approach of bad debt per household).
- We carried out a fresh assessment of each of the explanatory variables used for cost drivers in the PR19 models, looking in particular at their performance in the updated dataset.
- In relation to variables used to capture cost drivers relating to deprivation and arrears risk, we explored alternative candidate variables – drawing primarily on the wider set of variables related to arrears risk included in the updated dataset from Ofwat – and alternative functional forms.
- We explored whether there were grounds for including explanatory variables that would allow for differences between the costs of retailing water services compared to the costs of retailing wastewater services to be captured in the models.

Approach to model assessment and review

We used a range of information, evidence and analyses to explore the relative merits of alternative model specifications and to compare these against Ofwat's models from the PR19 review.

We summarise in the diagram below the aspects that we consider to be particularly relevant for the purposes of comparing the performance of alternative models or modelling approaches.

Figure 1 Dimensions of model performance and robustness we considered



The considerations above relate to a mix of qualitative and quantitative assessment. In terms of the quantitative assessment and statistical analysis, our approach used a combination of the following:

- **Established techniques**. We drew in part on the types of established metrics and tests that Ofwat has used in the past and which it told companies to report on for the purpose of submitting models to it in January 2023. For instance, we gave particular weight to: the consistency of estimation results with prior expectations; the reported statistical significance (and t-statistics) of the estimated coefficients for explanatory variables and, where legitimate to do so, the R-squared measure of goodness of fit.

- **Enhanced metrics**. We also drew on some more sophisticated metrics and forms of quantitative analysis, which Reckon has developed as means to tackle some of the limitations of the metrics and tests used by Ofwat, and to bring a more reliable and insightful evidence base for the purposes of developing and selecting water company benchmarking models. The enhanced metrics relate closely to factors that Ofwat rightly cares about in selecting econometric models (e.g. goodness of fit and the sensitivity of estimation results to changes in input data).

Across all aspects of our model development and assessment work, we sought to take a proportionate and targeted approach within the time and resource available, giving greater attention to the evidence and issues that we considered to be most important.

The main outputs from our model development process are as follows:

- Insight and evidence on econometric modelling approaches that might be used for benchmarking residential retail activities at PR24.

- A suite of econometric models for residential retail activities, estimated using updated data, which we consider to be an improvement on the set of models used by Ofwat at PR19.

In preparing a suite of models of residential retail costs to present in the report, we decided to err on the side of having a larger rather than smaller number of models. This was for two main reasons. First and foremost, we see benefits from using a diversity of reasonable modelling approaches, rather than trying to select between models on what might be quite marginal or subjective evidence. In addition, we recognise that Ofwat and the client water companies will have their own views on what aspects of model performance, tests or metrics are most important and a broader suite of models allows for users to select a subset that better fits with their preferences and views.

As part of our model review process, we considered the set of tests and metrics that Ofwat asked water companies to report on for the purpose of econometric model submissions in January 2023. There was considerable overlap between the factors which Ofwat said it would attach high priority to and those which we gave weight to as part of our model development process. For one of Ofwat's medium priority tests, the RESET test, we found that not all the models we put forward had a high p-value, but we did not think that this pointed to any major concerns with these models, or grounds to narrow down the suite of models. It may however provide information to help guide further model development work in the future.

Aggregated versus disaggregated models

Ofwat's PR19 retail model suite included disaggregated models and aggregated models. We show in the diagram below the more granular elements of the residential retail costs reported by water companies which form part of total residential retail costs and how these are allocated between the two more granular categories used by Ofwat at PR19: bad debt related costs and other retail costs.

For our modelling we used the same breakdown as Ofwat used at PR19. We think it would be better if Ofwat asked companies allocate depreciation, amortisation and recharges between the two modelling categories of (a) bad debt related costs and (b) other retail costs. But data reported on this basis were not available to us and making an approximate allocation was not something we prioritised for the project.

Figure 2 Breakdown of aggregated and disaggregated modelling categories

Total residential retail costs
Debt management
Charges for bad and doubtful debt
Customer service
Meter reading
Other operating expenditure
Depreciation and amortisation
Net recharges

Bad debt related costs
Debt management
Charges for bad and doubtful debt

Other residential retail costs
Customer service
Meter reading
Other operating expenditure
Depreciation and amortisation
Net recharges

By aggregated models, we mean models where the dependent variables is defined with references to the measure of total residential retail costs from the figure above. By disaggregated models we mean models that use the measure of bad debt related costs in the dependent variable or models that use the measure of other residential retail costs in the dependent variable.

We consider it good practice to explore models at different levels of aggregation as part of the model development and review process. But this does not necessarily mean that the final suite of models should involve models at both levels of aggregation.

There are theoretical benefits and drawbacks of more aggregated models, versus disaggregated models, but we feel that, in the context of residential retail cost modelling,, the case in favour of aggregated models is at risk of being over-stated. This is especially so in a context where the aggregated models used in Ofwat’s wholesale and retail benchmarking tend to have quite visible disadvantages. For instance, at PR19 the t-statistics on the estimated coefficients in the aggregate cost models tended to be lower and those aggregated models suffered from the omission of cost driver variables that made sense intuitively and that performed sufficiently well in the disaggregated models.

The enhanced metrics of model performance that we have developed allow us to make valid comparisons of statistical performance between: (i) cost benchmarks derived from aggregated models; and (ii) cost benchmarks built up by summing modelled costs from disaggregated models.

Based on the analysis described in more detail in section 6 of this report, we consider there is empirical evidence that the aggregated models we explored are inferior in a key dimension of performance compared to the disaggregated models. In particular, compared to an approach of summing cost benchmarks built up from disaggregated models, the aggregated models seem to provide cost benchmarks that are considerably more sensitive to variations in the dataset (e.g. dropping companies and/or years), without a significant benefit in terms of goodness of fit.

While the enhanced metrics we have used are one way to compare the performance of disaggregated models and aggregated models, it is also possible to use Ofwat's established techniques to see some of the limitations of the aggregated models. For instance, we found a pattern that, for the same set of explanatory variables, their estimated t-statistics were (in almost all cases) lower, and hence their statistical significance worse, when estimated within an aggregated model compared to when estimating disaggregated models.

In light of the analysis presented in this report, we decided that our suite of econometric benchmarking models for residential retail costs should focus on disaggregated models, which take bad debt related costs and other retail costs separately, and to not include models of total residential retail costs. We take each of these two disaggregated cost categories in turn below and summarise our findings.

Models of bad debt related costs

In line with Ofwat's approach at PR19, we developed and estimated models of bad debt related costs. (with the scope of bad debt costs as defined above).

We briefly summarise some of the main findings from our work to explore alternative specifications for econometric models of bad debt costs:

- **Dependent variable**. We found a good case for including models in which the dependent variable is specified as bad debt related costs divided by total billed revenue, *alongside* models where the dependent variable is specified in terms of bad debt related costs per household (as used for PR19). The former is a more restrictive model specification, imposing an assumption of a 1:1 relationship between a company's bad debt related costs and the average size of its bills, unless average bill is included as an explanatory variable. But that assumption seems quite a reasonable one from an economic perspective, *at least as an approximation*. Given the small sample size and complexities of companies' cost structures, there can be benefits from imposing simplifying assumptions rather than allowing the relationship to be estimated by the model. Our analysis indicates that the results for models involving this assumption were considerably less sensitive to variations in the dataset than corresponding models where the dependent variable was expressed on a cost per household basis, with some evidence of slightly improved goodness of fit too.
- **Cost driver relating to deprivation and credit risk (1)**. As at PR19, there seems to be a role for (a) models which include an explanatory variable based on measures for economic

deprivation from the ONS and Statistics Wales, and (b) models that use an explanatory variable based on measures of customer credit risk derived Equifax data. We investigated the possibility of producing combined models which incorporated both of these data sources in the same model, specified in a way to reduce the extent of correlation between them, but this exercise was not successful in the time we had available for it.

- **Cost driver relating to deprivation and credit risk (2)**. We analysed the three different Equifax variables included in the dataset published by Ofwat. Of those, the variable RGC102 seemed to work best statistically and we consider that, intuitively, it is also the most appealing one to draw on for the modelling. This is a different Equifax variable to that used by Ofwat in its PR19 final determination models but is the Equifax variable that we had given emphasis to in a previous project for United Utilities at PR19.
- **Cost driver relating to deprivation and credit risk (3)**. There are some intuitive and statistical grounds for considering models in which the ONS/Statistics Wales deprivation variable is in the form of squared terms, calculated as the company-level aggregate of the square of the LSOA-level income deprivation score. This is to capture for non-linear and convex relationship between LSOA-level deprivation and bad debt relation costs.
- **Customer transience cost driver**. The explanatory variable from the PR19 models that was used to proxy for customer transience did not seem stable across model specifications and dataset variations, and sometimes gave counterintuitive results. While there is some logic for a customer transience cost driver, the specific variable used at PR19 to capture this has significant limitations and we did not consider that it worked reliably in the models we estimated.
- **Dynamic aspects of model specification**. We explored a number of different ways to specify dynamic aspects of the econometric models of bad debt costs. We found that the cost profile for bad debt costs over our sample period – 2013/14 to 2021/22 – did not fit well with either a constant term or a simple time trend. Our preferred models have either: (a) a time trend and a dummy variable for each of the three years since the start of the Covid-19 pandemic or: (b) a full set of year-specific dummy variables.

We set out below the 12 models of bad debt related costs that we included in the suite of models produced for this project.

We report these models over two tables, Table 1 and Table 2. The first of these table present the six models (models labelled D1 through to D6) where the dependent variable is expressed in terms of cost per household; whilst the second table reports on the six models (labelled D7 through to D12) where the dependent variable is expressed in terms of the ratio of bad debt related costs to revenue. All models are panel data random effects models.

In each of the two tables, we present the estimated coefficient for the relevant explanatory variables included in the model and, below each of these and within brackets, the value of the t-statistic. The table also shows, in the row labelled “Dynamics”, the dynamic specification assumed for the model: “Dummy vars.” Indicates that the model included a set of year-specific dummy variables, and “Trend plus 2020-2022 dummies” indicates that the model included a time trend plus a set of dummy variables for each of the three years in the period from 2019/20 to 2021/22. We have not included in the table the estimated coefficients on the year-specific dummy variables.

We follow Ofwat's convention of using asterisks to flag the statistical significance of the estimated coefficients: * indicates significance at the 10% level; ** at the 5% and *** at the 1% level.

Table 1 Models of bad debt related costs (1)

Model Ref	D1	D2	D3	D4	D5	D6
Dependent variable	Ln(Bad debt related costs per household, CPIH adjusted)					
Explanatory variables						
Ln (revenue per household)	0.971*** (9.569)	0.999*** (10.158)	0.953*** (9.645)	0.975*** (10.564)	1.000*** (11.050)	0.959*** (10.907)
Credit risk score (eq_rgc102)		-0.031 (-1.590)			-0.029** (-2.033)	
Income deprivation score (unadjusted)	0.046* (1.671)			0.046* (1.7)		
Squared income deprivation score (unadjusted)			13.395* (1.871)			13.286* (1.880)
Time trend				-0.045*** (-2.701)	-0.047*** (-2.733)	-0.046*** (-2.744)
Dynamics	Dummy vars.	Dummy vars.	Dummy vars.	Trend plus 2020–2022 dummies	Trend plus 2020-2022 dummies	Trend plus 2020–2022 dummies
Overall R-squared	0.74	0.73	0.73	0.74	0.73	0.72
Observations	153	153	153	153	153	153

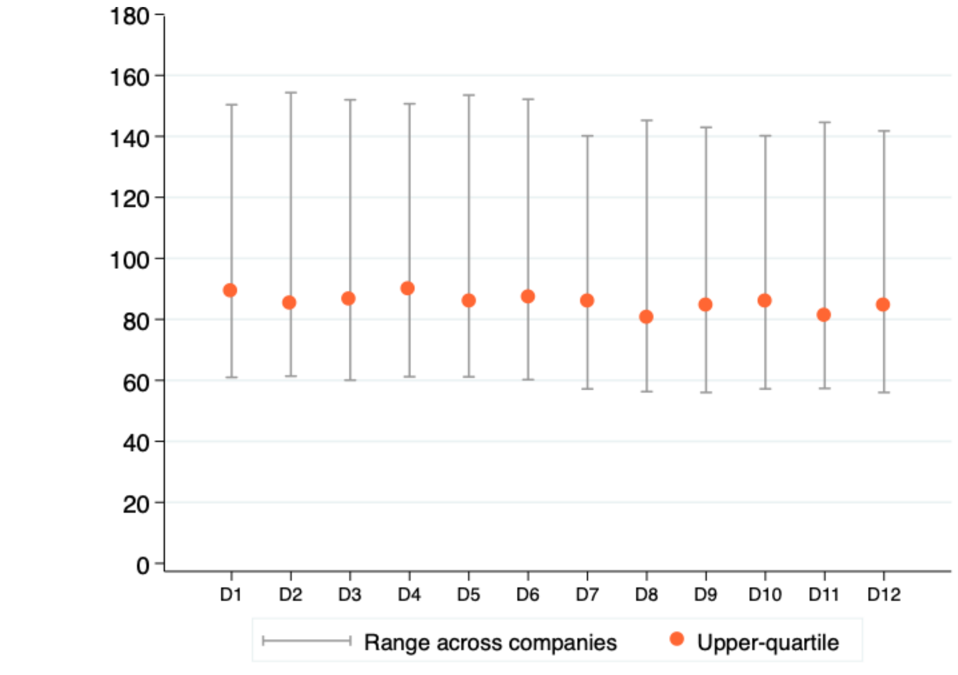
Table 2 Models of bad debt related costs (2)

Model Ref	D7	D8	D9	D10	D11	D12
Dependent variable	Ratio of bad debt related costs to total billed revenue					
Explanatory variables						
Credit risk score (eq_rgc102)		-0.001 (-1.427)			-0.001** (-2.010)	

Model Ref	D7	D8	D9	D10	D11	D12
Income deprivation score (unadjusted)	0.002* (1.887)			0.002* (1.916)		
Squared income deprivation score (unadjusted)			0.434* (1.755)			0.434* (1.778)
Time trend				-0.002*** (-3.378)	-0.002*** (-3.403)	-0.002*** (-3.394)
Dynamics	Dummy vars.	Dummy vars.	Dummy vars.	Trend plus 2020–2022 dummies	Trend plus 2020-2022 dummies	Trend plus 2020–2022 dummies
Overall R-squared	0.24	0.21	0.21	0.23	0.21	0.21
Observations	153	153	153	153	153	153

We have calculated the efficiency scores for each of the models set out in the tables above. These are based on comparing, for each company, actual and modelled costs across the last five years of data, from 2017/18 to 2021/22. Figure 3 shows the results. The grey line indicates the spread of the scores across companies – from the minimum to the maximum efficiency ratio – and the marker indicates the position of the upper-quartile efficiency ratio.

Figure 3 Efficiency ratios for set of bad debt related cost models



Models of other retail costs

We now tun to the models of other retail costs. These involved benchmarking the element of each company's total residential retail costs that remains once bad debt related costs (as defined above) are excluded.

We briefly summarise some of the main findings from our work to explore alternative specifications for other retail costs:

- **Dependent variable**. The main types of models we estimated involved the dependent variable being specified as the natural logarithm of other retail costs per household, as used by Ofwat for corresponding models at PR19. We also saw a role for models in which the dependent variable is defined as the natural logarithm of cost per service supplied. Statistically, this second approach seemed to perform less well than models where the dependent variable is expressed in terms of cost per household, but it has some intuitive appeal and avoids a potential concern we found when benchmarking costs on a per household basis (see point below on the proportion of dual service customers).
- **Economies of scale cost driver variable**. Some of the statistical evidence, as well as regulatory policy considerations, might suggest not including in the model specification a variable that allows for company-level economies of scale, as was done in some of the PR19 models. There is an argument that smaller companies have opportunities to mitigate their smaller size (e.g. through joint ventures or commercial agreements with other companies billing some of the same customers). However, it is also possible that smaller companies may not be able to fully offset scale effects (perhaps due to relative bargaining power) and some of the statistical evidence we collated supported the inclusion of an economies of scale explanatory variable in the case of models expressed on a cost per household basis. We followed the PR19 approach of including some models with an explanatory variable for economies and some without.
- **Meter penetration cost driver variable**. As for Ofwat's PR19 models, we found intuitive and statistical support for the inclusion of an explanatory variable for meter penetration, to pick up the additional costs arising from a greater proportion of the customer base being metered (whether for meter readings or dealing with customer contacts relating to metered charges).
- **Proportion of dual service customers**. For those models in which the dependent variable is defined in terms of cost per household, we found intuitive and statistical grounds for including an explanatory variable to capture the proportion of dual service customers. Indeed, we do not think that models to benchmark cost per household make sense without this explanatory variable. The costs of a retailer providing a single household with water and wastewater services will be greater than if the retailer were providing the household with only a water service or only a wastewater service. However, we did identify some potential concerns with the scale of the coefficient on this variable, which is one of the reasons for also considering models expressed on a cost per service basis.
- **The relative costs of retailing water and wastewater services.** In terms of potential cost drivers not captured in the PR19 models, we agreed with the client companies to focus on the potential for there to be differences in costs between retailing water services and retailing

wastewater services. These might relate to underlying differences in the retail activities for each (e.g. the subject matter of customer enquiries) or to differences in bargaining power between water-only companies versus water and wastewater companies. In the context of the models of costs per service, we found support for an explanatory variable to allow for differences in retail costs between water and wastewater retailing. This suggested that wastewater retailing might be lower cost on average.

- **Dynamic aspects of model specification.** We found that models that allowed for no dynamics, as those used at PR19, did not look appropriate on the latest data, which showed significant real-term cost reductions over the sample period. However, there does not seem to be the same degree of need to control for Covid-19 related factors as we found for the bad debt models. Our preferred models have either: (a) a simple time trend; or (b) a full set of time dummies.

We set out below in Table 3 the six models of other retail costs that we included in the suite of models produced for this project. The table is structured in the same way as the model result tables presented earlier.

The first four models, models O1 through to O4, are those where the dependent variable is expressed in terms of other retail cost per household, and models O5 and O6 are ones where the dependent variable is expressed in terms of cost per service. All six models are estimated as panel data random effects models.

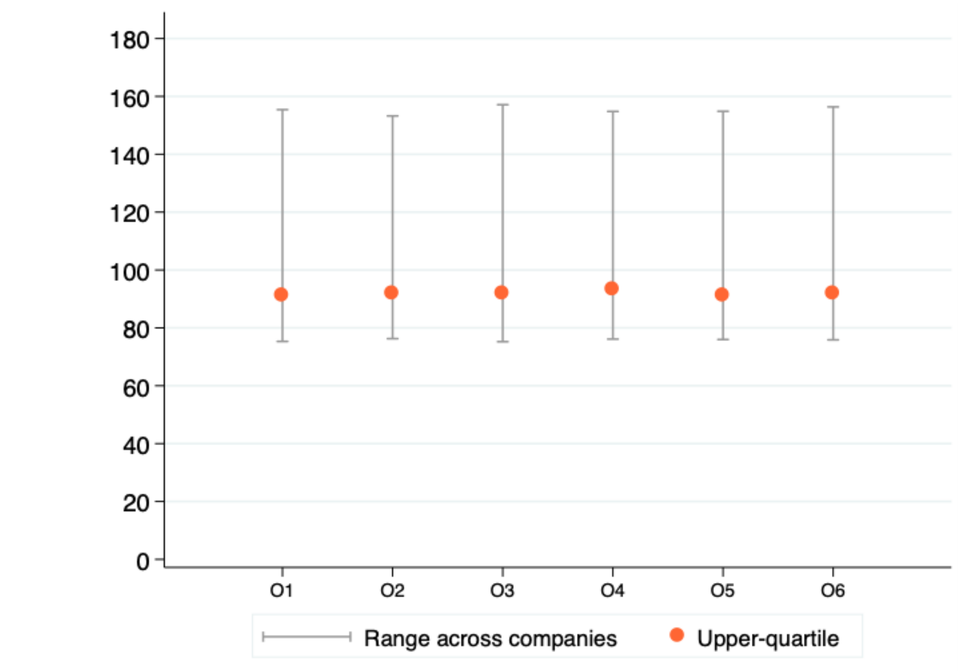
Table 3 Models of other retail costs

Model Ref	O1	O2	O3	O4	O5	O6
Dependent variable	Ln (Other retail costs per household, CPIH adjusted)				Ln (Other retail costs per service, CPIH adjusted)	
Explanatory variables						
Proportion of dual service customers	0.208** (2.367)	0.286*** (3.597)	0.21** (2.395)	0.292*** (3.562)		
Proportion of metered connections	0.495* (1.709)	0.509* (1.798)	0.555** (2.045)	0.577** (2.163)		
Ln (Total households connected)		-0.033 (-1.053)		-0.035 (-1.095)		
Prop. of measured services					0.603* (1.938)	0.663** (2.244)

Model Ref	O1	O2	O3	O4	O5	O6
Prop. of services that are wastewater					-0.717*** (-4.431)	-0.718*** (-4.558)
Time trend			-0.020* (-1.945)	-0.020* (-1.951)		-0.021** (-2.093)
Dynamics	Dummy vars.	Dummy vars.	Trend	Trend	Dummy vars.	Trend
Overall R-squared	0.18	0.20	0.15	0.16	0.48	0.46
Observations	153	153	153	153	153	153

Figure 4 shows the range of efficiency scores across water companies – from the minimum value to the maximum value, indicated, for a given model, by the grey vertical line – and shows the position of the upper-quartile company.

Figure 4 Efficiency ratios for set of other retail cost models



2: Approach to developing and reviewing models

This section gives an overview of the approach we took to developing and reviewing models – and modelling approaches – for the purposes of residential retail cost assessment. It takes each of the following points in turn:

- Building on a detailed review of Ofwat's PR19 econometric models.
- Models and modelling approaches.
- High-level outline of model assessment criteria.
- Enhanced metrics and analytical tools to support model assessment.
- Interactions with tests and metrics from Ofwat model submission guidance

Building on a detailed review of Ofwat's PR19 econometric models

The use of econometric benchmarking models for price control cost assessment is an ongoing, evolving, process, through which improvements can be made over time. There is not necessarily a target endpoint of a "robust" set of models, or a "correct" modelling approach. But it is reasonable for water companies and other stakeholders to expect significant improvements from one price control to the next.

Our approach has been to take the set econometric models for residential retail costs that Ofwat used in its PR19 final determinations as a starting point and to carry out a detailed review of these models, as a means to both understand potential limitations and drawbacks of these models and to help identify practical opportunities for improvements.

It is almost always possible to improve upon the existing set of models with sufficient expertise, data and time. To achieve good progress requires judgement on which areas to focus on (e.g. which limitations of current models to try to tackle) and creative thinking to identify new ways to specify and assess models.

We have sought to take a proportionate approach in this project, looking to explore and take opportunities for improvement against the PR19 models without seeking to exhaust all reasonable lines of potential model development or refinement.

Drawing on our review of Ofwat's econometric models at PR19, we agreed with Bristol Water and Wessex Water which areas of model specification, and which ideas for potential improvement, to focus on when exploring alternative models.

We have also drawn on insights we have developed through work on previous assignments for other clients concerned with benchmarking residential retail costs in particular.

During the course of our work, in November 2022, Ofwat published a dataset with information on costs and cost drivers relevant to residential retail benchmarking, and our analysis has drawn on that. During the period, Ofwat also published its PR24 Final Methodology and we have sought to reflect in our approach on the positions set out there, where relevant.

Models and modelling approaches

In looking to improve upon the PR19 models, our view is that focus is needed on what we call modelling approaches (defined very broadly) rather than specific model specifications. We have given attention to alternative modelling approaches that could be used for PR24 which seem likely to bring improvements relative to the modelling approaches embodied within the PR19 models.

We use the term modelling approaches in quite a broad way here, and it might include for example:

- the inclusion of a specific explanatory variable that does a better job of capturing an underlying cost driver than in previous models – or which captures a cost driver that has so far been overlooked in the modelling;
- a specific way of incorporating an explanatory variable within an econometric model (e.g. the logical arguments around whether it should be in logs or the case for using the square of a variable in the model);
- the way that the dependent variable is specified (choice between using aggregate costs or cost per customer in the dependent variable);
- the way that dynamics and changes over time in costs or cost driver relationships are to be accommodated in the model specification (e.g. a constant term, time dummy variables or a time trend).
- the type of model estimation technique (e.g. pooled OLS or random effects estimated using maximum likelihood or GLS)

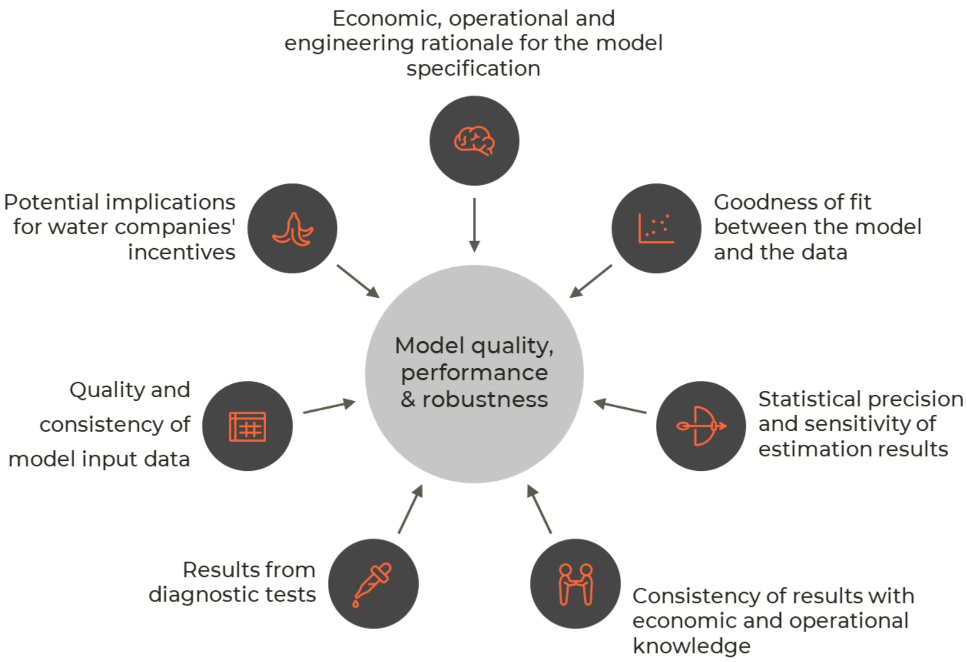
High-level outline of assessment criteria

We used a range of information, evidence and analysis to explore the relative merits of alternative model specifications and to compare these against Ofwat's models from the PR19 review.

We summarise in Figure 5 the considerations that we see as particularly relevant for the purposes of comparing the performance of alternative models or modelling approaches.

There is considerable degree of overlap between the set of dimensions presented in the figure above and (a) the set of PR24 base cost assessment principles that Ofwat's PR24 final methodology (insofar as these are relevant to the econometric modelling) and (b) the considerations lying behind the tests and metrics that Ofwat told companies to report on for the purpose of model submissions in January 2023.

Figure 5 Dimensions of model performance and robustness that we have considered



In terms of the quantitative assessment and statistical analysis falling under the dimensions above, our approach used a combination of the following:

- **Established techniques**. We drew in part on the types of established metrics and tests that Ofwat has used in the past and which it told companies to report on for the purpose of model submissions in January 2023. For instance, we gave particular weight to: the consistency of estimation results with prior expectations; the reported statistical significance (and t-statistics) of the coefficients for explanatory variables and, where legitimate to do so, the R-squared measure of goodness of fit.

- **Enhanced metrics**. We also drew on some more sophisticated metrics and forms of quantitative analysis, which Reckon has developed over time to as means to tackle some of the limitations of the metrics and tests used by Ofwat in its published work, and to bring a more reliable and insightful evidence base for the purposes of developing and selecting water company benchmarking models. The enhanced metrics relate closely to factors that Ofwat rightly cares about in selecting econometric models (e.g. goodness of fit and the sensitivity of estimation results to changes in input data).

We provide more details in the table below on the types of information we have drawn on, for each of the dimensions from Figure 5. The subsequent subsection briefly introduces the enhanced metrics and tools that we have used (further details of these are provided in appendix 1).

Table 4 Metrics and factors relevant to considering dimensions of assessment

Dimension	Brief summary or comment
Economic and business rationale for the model specification	We consider the economic and business rationale for alternative model specifications and modelling approaches, including the choice of candidate cost drivers and the way that these are incorporated as explanatory variables within specific models. This aspect of assessment involves looking closely at model specifications, before considering any regression results.
Goodness of fit between the model and the data	The goodness of fit concerns the extent to which the variation in the dependent variable (e.g. variation in unit costs between companies and over time) that is observed in the data can be explained by the model. It is related to the scale of residuals. There are alternative metrics to capture the goodness of fit, and we draw on several in our review. Specifically: • R-squared statistic. • Spread of implied efficiency ratios. • Normalised root mean square error. The last of these metrics is one we have developed for the purpose of better comparing goodness of fit across models for cost assessment benchmarking. We discuss this later in this section.
Statistical precision of estimated cost benchmarks	The ultimate role of the econometric models is to produce cost benchmarks for setting price controls, based on the predicted values from the estimation of the econometric models. The statistical precision of those predicted values is a relevant aspect to take into account. We review the precision of estimated cost benchmarks through a number of routes: • t-statistics and estimated variance for estimated coefficients. • Estimates of the variance of the predicted unit cost benchmarks produced by models and modelling approaches under consideration. This provides an indicator of the statistical sensitivity of these predicted values, having imposed the model on the dataset. We have considered this variance of predicted values using both the conventional variance estimates reported by Stata and also using estimates of variance derived using sensitivity of predicted values to changes in the data used. We discuss this metric later in this section.
Consistency of results (e.g. sign and magnitude of coefficients) with economic and operational knowledge.	It is relevant to review the sign and magnitude of estimated coefficients to check their consistency with what we would expect (insofar as there are prior expectations) given the identified economic and business rationale for the cost driver.
Diagnostic tests	It can be helpful to carry out certain “diagnostic tests” on econometric models, drawing on the types of tests used by Ofwat during PR19. We do not consider that diagnostic tests should be treated as the key pass/fail criteria for whether models are robust, and their importance to the overall model development process is sometimes overestimated. But we recognise that they have a role to play – not least in revealing aspects of model results that should be investigated further and in anticipating potential criticisms of models that might be put forward.
Quality and consistency of input data used by the model	The robustness of econometric models is in part a function of the data feeding into them. Especially with a small data sample, it is not sufficient to rely on any problems with input data being detected in the modelling results or statistical tests.

Dimension	Brief summary or comment
	We consider data quality as part of the review of models and modelling approaches.
Potential implications for companies' incentives	While not a direct feature of the economic and statistical performance of a benchmarking model, it is important to also consider the risk that the use of a specific model (or variable) in setting allowances at PR24 might adversely affect companies' incentives for efficiency or otherwise distort their behaviour. It is possible that Ofwat could rule out an otherwise good model on these grounds, so it is helpful to consider it as part of the development and review of models. One example where Ofwat might be concerned is the use of explanatory variables which are under management control to a significant degree, for which the use might dampen incentives for efficient operations. We keep this issue in mind in exploring options for model specification.

Across all aspects of our model development and assessment work, we sought to take a proportionate and targeted approach within the time and resource available, giving greater attention to the evidence and issues that we considered to be most important.

Enhanced metrics and techniques to support model assessment

For this project we have drawn on several enhanced metrics and analytical techniques that we consider to be particularly useful for the development and review of econometric benchmarking models, but which have not so far formed an established part of the toolkit used by Ofwat.

These build on the conceptual foundation provided by the set of criteria from Figure 1 and the types of considerations that Ofwat rightly cares about They are designed to overcome some of the limitations of existing approaches, and to give a richer and clearer picture of the ways in which potential new model specifications might provide improvements on current models – as well as exposing some of the trade-offs that can sometimes be faced in model development and selection. These metrics use computational and statistical techniques that are more advanced than those routinely used in water industry cost benchmarking.

We briefly introduce the main metrics and techniques:

- **Analysis of the sensitivity of estimated coefficients to variations in dataset**. One technique we draw on is to systematically examine the sensitivity of estimated coefficients to variations in the dataset used for the estimation (variations created by dropping a single year and/or a single company from the dataset). This is particularly useful when looking to assess whether specific explanatory variables are giving results that might be spurious or misleading, especially in the context of econometric models estimated on a relatively small data sample. A histogram of estimated coefficients over runs of the same model over a large number of dataset variations is a good way to help gauge the sensitivity of estimated coefficients. This can complement the more established approach of looking at purported statistical significance and t-statistics.

- **Tailored metric of goodness of fit**. Goodness of fit is an important aspect of model performance. At PR19, the main metric used for this by Ofwat is R-squared, but that metric suffers from major limitations. In particular, it is not directly comparable across models with different dependent variables and inferences on model performance based on R-squared can be misleading. We have maintained an emphasis on goodness of fit for comparing the performance of models, but used an alternative metric which is designed to be comparable across models with different dependent variables (and which can also be used to compare different suites, combinations or aggregations of models against each other). This metric is based on the established statistical concept of root mean square error (RMSE) as a measure of goodness of fit between a model to observed data. We have calculated a version of RMSE which captures the relative divergence between actual and modelled costs.

- **Tailored metric of the statistical sensitivity of estimated cost benchmarks.** The ultimate role of the econometric benchmarking models is to produce cost benchmarks for each company to be used in setting price controls, based on the predicted values from the estimation of the econometric models. In this context, we consider that it is important – and illuminating – to consider the statistical variation or sensitivity of the predicted values estimated by each model. At the simplest level, one can calculate statistical confidence intervals around the predicted values (and hence cost benchmarks) and these will be a function of the a confidence intervals around each of the individual coefficients in the model. This is something that seems to have been given relatively limited attention to date in Ofwat's econometric previous modelling of water company costs, and in its recent guidance on model submissions for PR24, but which seems highly relevant given the way models are used. Our approach has been to use a measure of how sensitive the costs predicted for each company are to minor variations in the dataset used (e.g. re-running the regression, dropping one company and/or year from the sample).

We provide a more detailed description of these metrics and techniques in appendix 1.

Interactions with tests and metrics from Ofwat model submission guidance

Ofwat invited water companies to submit potential econometric models for PR24 to it in early January 2023, ahead of a planned Ofwat consultation in spring 2023.

As part of our model review process, we considered the set of tests and metrics that Ofwat asked water companies to report on for the purpose of the model submissions in January 2023. There was considerable overlap between factors to which Ofwat said it would attach high priority and those we had already given weight to as part of our model development process.

For other test and metrics identified by Ofwat, we looked at the results (where applicable to the modelling approaches we were using) and considered what implications, if any, they should have for the selection of models for the suite presented in this report. For instance, we considered whether we should narrow down the set of models we had provisionally identified, in light of the difference in how different models performed in terms of these tests and metrics.

As Ofwat's guidance seems to recognise, in some cases the tests do not necessarily imply any problems with a model under consideration but may provide information to help guide further model development work to try to identify improved models.

3: Review of Ofwat's PR19 residential retail benchmarking

In this section we present our review of Ofwat's PR19 models – and modelling approaches – for residential retail activities. We used this as a means to both understand potential limitations and drawbacks of these models and to help identify practical opportunities for improvements.

We discuss the following aspects of Ofwat's PR19 models in turn:

- PR19 model suite and key estimation results reported.
- Use of aggregated and disaggregated models.
- Coverage of cost drivers.
- Estimates of cost driver relationships.
- Model dynamics.
- Random effects versus pooled OLS.
- Smoothing of depreciation data.
- Allocation of depreciation and recharges.

For each of these aspects, we explain how our review of the PR19 models and approach has influenced the direction of the modelling work that we carried out during the project.

The final subsection in this section summarises the specific areas of model specification, and ideas for potential model improvement, that we agreed to focus on with Bristol Water and Wessex Water to prioritise when exploring alternative models for the purposes of the project. Our identification of opportunities and priorities reflected the outcome of our review of the PR19 models.

In this report we have not sought to compare Ofwat's cost allowances with outturn data and have focused instead on a review that is intended to be without the benefit of hindsight.

PR19 model suite and key estimation results reported

In the table below we provide a copy of the table of model specifications for residential retail costs that Ofwat used at PR19 and the key estimation results that Ofwat presented alongside these model specifications. The asterisks by the coefficients indicate what Ofwat reports as statistical significance at the 99% (three asterisks), 90% (two asterisks) and 80% confidence levels (one asterisk). Where no asterisk is reported Ofwat provided the t-statistic.

Figure 6 Model specifications and key estimation results reported by Ofwat in its PR19

Model name	RDC1	RDC2	ROC1	ROC2	RTC1	RTC2	RTC3
Dependent variable (log)	Bad debt and bad debt management costs per household		Other retail costs per household		Total retail costs per household		
Average bill size (log)	1.190***	1.158***			0.458***	0.526***	0.603***
Proportion of households with default (Equifax variable) (%)	0.067***				0.024 {0.106}	0.030**	
Proportion of households income deprived (income score of IMD) (%)		0.076***					0.059***
Total migration (% of population)		0.035**					0.037**
Proportion of dual service households (%)			0.002*	0.002**			
Proportion of metered households (%)			0.007***	0.007***	0.004 {0.321}	0.004 {0.206}	0.002 {0.436}
Number of households connections (log)				-0.039 {0.394}		-0.059*	-0.116**
Constant term	-6.032***	-5.680***	2.400***	2.909***	-0.014 (0.980)	0.226 (0.653)	0.200 (0.564)
Overall R-Squared	0.77	0.78	0.13	0.15	0.67	0.70	0.71
Number of observations	105	105	105	105	105	105	105

We provide comments on the choice of explanatory variables, and the estimated coefficients for some of these variables, further below. Here, we make some more general comments, especially in terms of the goodness of fit of the models:

- The two models focused on bad debt costs perform reasonably well in terms of R-squared (0.77 and 0.78), especially given that these are models of unit costs (cost per household) rather than models of aggregate costs as Ofwat used for its wholesale models. For aggregate cost models, it tends to be straightforward to achieve a high R-squared due to the explanatory power of the relevant scale variable, but we generally expect a lower R-squared for unit cost models.

- Despite the favourable R-squared of the bad debt models, the range of efficiency scores (calculated over the last five years of data) is quite high in these models: 0.64 to 1.51 across the two bad debt models if these are weighted equally. This suggests that the input data on the dependent variable – the natural logarithm of bad debt related costs per customer – show a high degree of variation across companies and that considerable variation remains which is not explained by the models despite these models explaining a large proportion of that variation.

- The R-squared of the two models for "Other retail costs per household" is very low (0.13 and 0.15), showing the explanatory variables explain only a small proportion of the variation across companies and over time in other retail costs per customer. But these models show a lower range of efficiency scores than the bad debt models: 0.68 to 1.37 (calculated over the last five years of data).

- The R-squared for the models of total residential retail costs lie between those for the bad debt models and other retail cost modes.

Looking across Ofwat's PR19 suite of models, the results suggest significant potential for improved models to be identified that could provide more robust results and better explain the variations seen in the data than these models.

Use of aggregated and disaggregated models

As indicated in the table above, Ofwat used a suite of models which included aggregate models, where the dependent variable related to total residential retail costs, as well as a set of more disaggregated models which took each of bad debt related costs and "other retail costs" separately. At least as a starting point, Ofwat's approach of estimating models at alternative levels of aggregation seems sensible.

Ofwat referred to the former as its top-down models and the second type as bottom up models. We feel that the term bottom-up models might be a little misleading for the second type and we refer to the two types as the aggregated models and disaggregated models.

There seem to be some clear drawbacks in Ofwat's aggregated models relative to the disaggregated models:

- The dual service explanatory variable (proportion of households to whom the retailer supplies both water and wastewater retail services) features in the disaggregated models relating to other retail costs (ROC1 and ROC2) but is absent from the aggregated models (RTC1, RTC2, RTC3). This variable has valid economic and business rationale as a cost driver in the context of models of retail costs *per household*. This means that the aggregated models miss out on a relevant cost driver that is taken into account in the disaggregated models. In the absence of the dual service variable, other explanatory variables in the aggregated model might pick up its effects in an indirect way, via correlations between the proportion of dual service customers and the explanatory variables in these models (e.g. average bill size and total number of customers supplied).
- For two of the explanatory variables that feature in disaggregated models as well as in aggregated models, the reported statistical significance of the estimated coefficients is lower/worse in the aggregated models. This applies to the proportion of households with default (Equifax variable) and, most dramatically, to the proportion of metered households which Ofwat reports as statistically significant at the 99% confidence level for both of the disaggregated models where it is included and not statistically significant at the 90% confidence level for any of the three aggregated models.

While the R-squared for each of the models of other retail costs is very low, it is not reasonable to compare this to the R-squared of an aggregated model as the dependent variable is completely different.

In setting allowances, Ofwat gave a 75% weight to results from its aggregated models and 25% weight to the results from its disaggregated models. This is in contrast to its approach for wholesale water and wholesale wastewater activities at PR19 where it gave equal weight to the aggregated models and the disaggregated models. Ofwat explained this as follows:

> *"Given the relatively wide spread of efficiency scores (ie residuals) in bad debt models compared with the total retail costs models, we give our bottom up view of costs a weight of 25%, with a higher weight of 75% to our top down view of costs.*
>
> *In response to Dŵr Cymru and SES Water's representations on the wide spread of efficiency scores, we consider that the spread of efficiency scores of the 'other costs' and total costs retail models is reasonable and similar to the spread of efficiency scores of our wholesale models. The spread of our retail models, including our bad debt models, is significantly lower than the spread of our standalone growth models from the initial assessment of plans. We acknowledge the wider spread of efficiency scores in bad debt cost models relative to the rest of our retail models. Consequently, we reduce the weight of the bottom-up models and instead put more weight on our top down models when calculating allowed residential retail costs."*

In the extract above, Ofwat points to "the wider spread of efficiency scores in bad debt cost models relative to the rest of our retail models". However, that by itself says nothing reliable about the relative quality of disaggregated models compared to aggregated models. It is not a like-for-like comparison. If comparing efficiency scores, what should matter is the efficiency score from the disaggregated models when results are summed up across bad debt models and other retail cost models versus the efficiency score from the aggregated models.

We have not checked Ofwat's calculations in detail but, using its own spreadsheet of reported efficiency scores, we found the range for the efficiency scores calculated from the set of four disaggregated models was 0.81 to 1.34, compared to a range of efficiency scores from the three aggregated models of 0.83 to 1.34. This does not seem a large difference at all and raises questions about the validity of Ofwat's reasoning for giving a higher weight to aggregated models.

Ofwat's approach here is concerning given that there are clear weaknesses in the aggregated models, including the omission of the dual service cost driver variable and greater estimation uncertainty, as highlighted above.

In its PR24 draft methodology, Ofwat said the following in relation to aggregated versus disaggregated models:

> *"The exception to this is our residential retail cost models, where we are proposing to simplify our approach by using the aggregated top-down model only. We believe this simplification is appropriate because the top-down models performed better than bottom-up models at PR19 and produced similar results. We will only consider using the bottom-up models where there is a demonstrable reason to do so."*

We found this to be strange and concerning, in part because we did not see the evidence in favour of the aggregated models at PR19 and in part because it seemed unnecessary and premature to take a position on this aspect of the retail modelling at the PR24 methodology stage.

Ofwat adopted a different position in its PR24 final methodology:[1]

> *"In our draft methodology, we proposed to only use the bottom-up models where there is a demonstrable reason to do so to assess residential retail costs. [and] instead focus on the top-down residential retail models. Five companies questioned this as they thought it was appropriate to consider different cost drivers for different elements of retail costs, which could be facilitated more with bottom-up models. To clarify, we recognise the benefit to keeping open the option of using models at different levels of cost aggregation to assess residential retail costs at PR24. We are also open to considering fresh ideas that could improve residential retail cost assessment. We will continue to review retail cost drivers, including the impact of deprivation on companies' bad debt costs, ahead of the spring 2023 modelling consultation."*

For the purposes of this project, we have considered both disaggregated and aggregated modelling approaches. We have placed particular emphasis on using new forms of empirical evidence to assess the relative merits of each of these approaches.

Coverage of cost drivers

We now turn to the set of cost drivers that are captured – to some degree – by the explanatory variables in Ofwat's PR19 econometric models for residential retail costs. In the table below we list the set of cost drivers that are captured in those models as well as some potential further cost drivers that were not. The list of captured drivers reflects those included in at least one model within Ofwat's final suite of models.

Table 5 Understanding of potential retail cost drivers (illustrative – not exhaustive)

Cost drivers captured by the PR19 models (to some degree)	Other possible cost drivers
• Number of households supplied • Proportion of dual service vs single service customers • Measures of income deprivation for customer base • Measures of arrears risk produced by credit reporting agency for customer base • Size of retail bills • Meter penetration • Measures of customer transiency • Economies of scale with respect to number of customers	• Mix between retail water and wastewater services • Aspects of wholesale and retail service quality • Factors affecting relative ease of meter reading • Digital literacy and usage • Regional wages and input prices (insofar as these affect local costs such as meter reading) • Number of customers who are "off-grid" for water or wastewater services

Looking across the suite of PR19 models, our view is that these do a reasonable job of including a series of relevant cost drivers. The PR19 models represented a major improvement on the approach at PR14 which was not based around econometric modelling and took a much more

[1] Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24 Appendix 9 – Setting expenditure allowances*, page 10.

limited account of the cost drivers above. None of the cost drivers captured in the PR19 models seem unreasonable to include.

Our list of other possible cost drivers draws on our knowledge and experience from previous econometric modelling work for water industry retail costs as well as a session with Pelican, Bristol Water and Wessex Water to discuss potential drivers missing from Ofwat's PR19 models.

We also reviewed the cost adjustment claims for retail costs at PR19 to see if these indicated cost drivers that were omitted from the models. This is a useful check but did not add to the list above: these claims did not show any significant gaps versus the table above. One exception concerns the claim by DCWW for bilingual service provision, but we consider this too specific an issue to consider capturing in the econometric modelling and is more relevant to a cost adjustment claim, if material, so we do not discuss further below.

Of the cost drivers that are not captured in the PR19 models, we took the following position for the purposes of the project:

- **Mix between water and wastewater services**. We saw the exploration of the cost driver relating to the mix between retail water and wastewater services as something to consider further during the modelling phase of this project. Some of Ofwat's models (those which include the dual service variable) allow for the retail costs to be lower *per service* if the retailer provides both water and wastewater to the same customer than if it provides a single service. But Ofwat's PR19 models make an implicit assumption that the costs of providing a single retail service to a customer are the same regardless of whether that service is water or wastewater. Our work for another client earlier in the PR24 process raised questions about this assumption and we were able to develop models that allowed for the retail costs of the two services to differ. However, this can be a difficult issue to address in practice, in part because of the correlation observed in the data between (a) the mix of retail services between water and wastewater and (b) the size of a retailer (in a context where some form of economies of scale may apply).
- **Aspects of wholesale and retail service quality**. There is a reasonable argument that wholesale and retail service quality can act as a cost driver for retail costs, as it may cost more to provide a higher-quality service. However, Ofwat has generally shown reluctance to include aspects of service quality in its econometric benchmarking models (though leakage might be argued to be an exception in the case of wholesale models). Furthermore, our own experience is that there are difficulties in adequately capturing variations in service quality within the econometric model specifications. We agreed with the client companies that trying to do so was not a priority. Nonetheless, if quality variables are to be omitted from the benchmarking models used for PR24, it remains important to recognise them as a potential cost driver and consider the implications of this for wider aspects of Ofwat's cost assessment at PR24. For instance, companies that look "inefficient" from the modelling alone might actually be efficient and just incur a higher cost to provide better levels of performance. Furthermore, and related to this, Ofwat's upper quartile efficiency adjustments might not be reasonable if they overlook the potential for apparently lower-cost companies to be providing worse service quality.
- **Digital literacy and usage**. If there were evidence of substantial differences in digital literacy, and usage of online services, across different parts of England and Wales, this might be relevant

to the modelling, especially for other retail costs. Our initial discussions with Pelican indicated that the costs of customer service and billing (e.g. postage) can be significantly affected by whether customers manage their accounts and receive bills online. However, we were not aware of good data sources and metrics for digital literacy or of usage of online services that would allow valid comparisons across water companies. Furthermore, even if its effect were significant it may not be large enough to get reasonable results from econometric models estimated on noisy data in a small sample size. We agreed with the client companies not to explore this cost driver in the modelling phase of the project.

- **Factors affecting relative ease of meter reading**. There are a number of factors that may affect the costs of meter reading, such as the geographical dispersion of customers or the level of traffic congestion in the local area. We do not rule out the possibility that econometric models could be developed which pick this up, but this seems likely to be challenging given the small scale of meter reading costs within other retail costs and the limitations of econometric modelling on small sample sizes. We agreed with the client companies not to explore this cost driver in the modelling phase of the project.

- **Regional wages and input prices**. We consider that, in principle, regional wages and other local-level input prices could be a cost driver for retail costs that leads to differences across companies. However, given that few retail staff truly need to be based within the same area of appointment as the wholesale company, this seems quite a limited issue for retail costs, perhaps limited primarily to meter reading. There is an argument that input prices could be relevant to explaining changes over time in costs, but we would imagine that Ofwat would be highly resistant to inclusion of input prices in the modelling. We agreed with the client companies not to explore this cost driver in the modelling phase of the project.

- **Number of customers who are "off-grid" for water or wastewater services**. This is a subtle point, and we are uncertain whether it is material. It is widely recognised that there are some economies of scope available when retail water services and retail wastewater services are supplied to the same residential customer. However, the opportunities for economies of scope do not depend on those customers being formally both retail services supplied by the same company. This is because in practice water companies make arrangements to benefit from economies of scope even for customers served by separate wholesales. The Pelican joint venture between Bristol Water and Wessex Water is one such arrangement. The more common arrangement is for one company to carry out billing and other retail activities on behalf of another company serving the same customers (e.g. a WoC billing wastewater services on behalf of a WaSC) in exchange for payment. However, if there are customers who are "off-grid" for one of the services (i.e. they do not take wastewater services from any company) then there would be no opportunity for economies of scope to be achieved for that customer and it would have a higher cost to serve. If some retailers had significantly more off-grid customers than others, it is possible that this could cause material differences in average costs per customer or per service. But we suspect that the small number of off-grid customers mean that it would be difficult to capture this type of cost driver in the econometric modelling. We agreed with the client companies not to explore this cost driver in the modelling phase of the project.

Estimates of cost driver relationships

Including a variable within an econometric model to capture a specific cost driver does not guarantee that the modelling will produce good estimates of the relationship between retail costs and that cost driver. In part this will depend on data quality, sample size, and how important the cost driver is in contributing to observed cost differences across companies. It will also depend on modelling choices such as the data source and variable used to capture the cost driver and other aspects of the model specification (e.g. level of aggregation, choice of dependent variable and other explanatory variables in the model).

In this section we make some targeted comments on two key aspects of the modelling results for the cost drivers included in the models:

- Statistical precision and sensitivity of estimation results.
- Consistency of results (e.g. sign and magnitude of coefficients) with economic and operational knowledge.

Our view is that, for the cost drivers captured by Ofwat's models, there are some limitations and concerns about the results from the PR19 modelling exercise and some scope for improvement at PR24. We highlight a few points below:

- **Dual service variable**. The inclusion in models of a variable capturing the proportion of dual service customers makes sense in the case where the dependent variable is retail cost per household, especially as a driver for other retail costs (excluding bad debt). However, in Ofwat's PR19 models, the coefficient for the dual service variable was lower than might be expected and is perhaps questionable. If a customer is not supplied on a dual service basis, they will often – though not always – be provided with *retail services* by the same company, whether through a joint venture or through agreements between the relevant company providing the water service and that providing the wastewater service. This allows for the costs to be split between the two companies, which we think should lead to the costs per customer for supplying retail water or wastewater separately being significantly less (and perhaps around half) of that for a dual service customer. But the estimated coefficients in the PR19 models imply less of a difference.
- **Economies of scale variable.** Given that the dependent variable is the cost per household, the inclusion of an explanatory variable for the number of households in some of Ofwat's models at PR19 provides a way to take account of economies of scale. The coefficient on this variable in those models is estimated to be to be negative and this acts to increase the estimate of efficient costs for companies that serve a small customer base relative to larger ones. However, the econometric results suggest that the estimated coefficient for that variable is not stable across models and has a low t-statistic for some models, which raises questions about model accuracy. Perhaps of more importance from a regulatory perspective, a question arises as to whether Ofwat should provide additional allowances to some companies due to their size, in the light of the potential opportunities for any actual cost inefficiency from smaller size to be mitigated at the retail level (e.g. via joint ventures or agreements with other companies).
- **Deprivation and arrears risk.** Ofwat's approach at PR19 benefitted from using two different ways to capture an underlying cost driver associated with customer's risk of falling behind or not paying their water bills, relevant for modelling bad debt related costs. Specifically, it drew on:

data on economic deprivation from the ONS and Statistics for Wales and data relating to arrears risk from Equifax.

- **Average bill size.** Average bill size makes sense as a driver of bad debt related costs and the econometric modelling seems to work reasonably well for this variable. However, the implication of the estimated coefficient is that, say, a 10% increase in average bills translates into a greater than 10% increase in bad debt costs; this is not altogether intuitive and is arguably more a reflection of estimation error/uncertainty for the coefficient on average bill size than an accurate reflection of the underlying cost structure.

- **Meter penetration.** It seems good to take account of meter penetration in the set of explanatory variables, as this will drive meter reading costs and is likely to affect customer service costs too. The main point to highlight from the PR19 results is that the statistical significance of the meter penetration coefficient drops considerably in moving from disaggregate to aggregated models.

- **Customer transiency variable**. We think that there are some questions about the specific variable used for customer transiency. While the variable used at PR19 was intended to capture customer transience or migration, it was not a direct measure of this. It was based on ONS data relating to migration flows in and out of local authority districts (LADs) which were used to construct a metric that proxied customer transiency. The ONS data on migration flows do not reflect household moves within an LAD, which could account for a significant share of population flows and changes to customers' addresses (e.g. students in Bristol and Bath moving from one address to another would tend not to be captured).

Model dynamics

None of Ofwat's PR19 models for residential retail allowed, within their specification, for the dynamic evolution of retail costs over the time period covered by the model input data: 2013/14 to 2018/19. This means that:

- The models make no allowance for changes over time in costs across the industry which are due to factors such as efficiency improvements or input price changes, beyond adjustments to deflate reported costs using CPIH.

- The models assume that the impact of any cost driver on costs is the same across the sample period (e.g. X% increase in costs for Y% change in the cost driver variable).

Of these, we consider that the first is the more concerning. Alternative approaches could be to include in the model specification a set of dummy variables for each year, or a variable capturing a time trend as well as more complex assumptions about the relationship between costs in successive years.

In addition, there was a limitation in terms of the availability of input data for a key explanatory variable used in the models of bad debt costs. For the ONS/Statistics Wales deprivation variable, values were available across all companies only for a single year and these were taken as the values across all years in the data sample (other than for companies in Wales for which more regular data was available). The effect of this would be to overlook any changes over time in deprivation and arrears risk in the customer basis, which could act to worsen the statistical precision of estimated coefficients for these variables and, in turn, wider modelling results.

Overall, Ofwat's PR19 models were highly static over its six-year dataset.

The question of including in the model specification a set of year-specific dummy variables seemed to have been considered, though perhaps only briefly, as part of the PR19 appeals to the CMA:[2]

> *"We have considered Professor Saal's suggestion of including time dummies in the model to capture changes that occur over time. While time dummies (either yearly or for full AMPs) may be helpful to capture common cost changes across the industry, they require strong assumptions when their coefficients are used for forecasting allowances. For example, time dummies would capture (at least partly) inefficiency too. For this reason, we decide not to use time dummies."*

We were a little surprised by this statement from the CMA which does not seem to stand up to scrutiny. To the extent that year-specific dummy variables might capture (at least partly) inefficiency, this is true also for including a constant term. Furthermore, the argument against time dummies that these require strong assumptions when used for forecasting allowances does not seem reasonable. Using time dummies might expose more explicitly the nature of assumptions being used to extrapolate from historical data to a forecast period. But strong assumptions are also needed under the approach used by the CMA and Ofwat where the model includes a constant term.

For the purposes of the modelling work for PR24, we agreed with the client companies to explore models with a range of different dynamic structures, covering time dummies, time trends only, and trend variables combined with time dummies. This more open approach to this aspect of the modelling seems particularly important for PR24 where the data period covers a period for which costs (especially bad debt costs) have been significantly affected by Covid-19.

Random effects versus pooled OLS

A further choice in terms of model specification and estimation technique is between random effects models (estimated using GLS or maximum likelihood) or pooled OLS. Pooled OLS has been used in the past but, over time, water industry cost benchmarking has gravitated towards a preference for random effects models, typically estimated using GLS.

Our experience is that, compared to other issues such as the choice of the cost driver explanatory variables and the functional forms of models, the choice between pooled OLS and random effects is of lower relative importance to benchmarking results (but it can still make a difference).

For the PR19 final determinations, Ofwat used only random effects models, explaining as follows:[3]

> *"Random effects models are appropriate for a panel data set (data with a cross section and a time dimension). The models also perform better statistically than under the*

---

[2] CMA (2021) *Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report*, page 201.

[3] Ofwat (2019) *PR19 final determinations: Securing cost efficiency technical appendix*, page 170.

ordinary least squares (OLS) method and the Breusch Pagan test consistently supports using random effects over the OLS."

Furthermore, for PR24 Ofwat has said the following:[4]

> *"We intend to use random effects estimation to estimate the base cost econometric models at PR24. This is widely supported by companies. Random effects estimation explicitly takes into account the panel data structure (ie repeated observations over time for multiple water companies), which is why it is preferred over standard ordinary least squares (OLS) estimation. We used random effects at PR19, and so did the CMA in the PR19 redeterminations. We would cautiously consider alternative estimation methods if random effects did not estimate sufficiently robust base cost models."*

Our view is that this seems a reasonable position.

It is conceivable that, in some cases, random effects models might not work well and pooled OLS could offer an improved model. But this issue did not seem a priority for our project and exploring both pooled OLS and random effects models is likely to be a distraction from more pressing matters.

Smoothing of depreciation data

We now turn to a few other aspects of the PR19 modelling approach which do not seem ideal.

At PR19, Ofwat smoothed depreciation over the six-year period spanned by the dataset in constructing the measure of deprecation to feed into the input data on retail costs for its econometric model.[5] Ofwat did not explain this in its final determination documents, but in a more detailed document on its econometric approach, Ofwat said the following:[6]

> *"In our 'other retail costs' and total retail cost models we include depreciation. Depreciation accounts for the reduction in the value of a company's fixed retail assets over time. The depreciation data in our sample is quite lumpy (being impacted by the lumpy nature of investment). We decided to smooth depreciation costs over the five-year period, to help reflect a more appropriate relationship of the cost with the explanatory factors in the model."*

We found this aspect of Ofwat's approach strange. The whole point of the accounting concept of depreciation is to smooth the expenditure incurred at a point in time over the anticipated period over which the benefits from that expenditure are realised. It is already smoothed. There may have been problems with the depreciation data available to Ofwat but its explanation of the nature of these or the rationale for its own smoothing does not make sense.

For our modelling work in this project we decided against using smoothed depreciation data.

---

[4] Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24: Appendix 9 – Setting expenditure allowances*, page 15.

[5] The calculation averaging depreciation over the six years covering the period 2013/14 to 2018/19 is set out in the worksheet "real_statafile" of the Excel file "FM_RR1_FD.xlsx" published by Ofwat at final determinations.

[6] Ofwat (2019) *Supplementary technical appendix: Econometric approach*, page 29. As noted earlier, the calculation carried out by Ofwat involved in fact averaging over six, and not five, years.

Allocation of depreciation and recharges

There is a further issue on depreciation to highlight. Ofwat's approach allocated all of the retail costs reported as depreciation to the set of costs modelled by the "other retail costs" model. Given this allocation, depreciation represents, on average across companies, around 9% of companies' other retail costs. Ofwat's full allocation of depreciation to the "other retail cost" category does not make complete sense from an economic and business perspective, since some of the assets (e.g. IT systems, office facilities) that are used for retail activities will be used in the course of debt management activities, so in principle we would expect total bad debt related costs – and the models that cover these costs – to include some element of depreciation. At best Ofwat's approach is an approximation.

The same issue applies to recharges. Ofwat's approach at PR19 allocated these to the "other retail costs" model; recharges account for around 5% of "other retail costs". But some of the costs reported as recharges will be costs that should be allocated to debt management activities (e.g. IT costs or facilities costs for staff working on debt management).

It would be more appropriate, given the sensible use of disaggregated models, for Ofwat to require that companies use appropriate internal data and assumptions to allocate their reported depreciation and recharges costs between debt management activities and other retail costs.

Priority areas for model development agreed with client companies

In this subsection we outline which specific areas of model specification, and ideas for potential model improvement, we agreed with Bristol Water and Wessex Water to prioritise when exploring alternative models for the purposes of the project. Our identification of opportunities and priorities reflected the outcome of our review of the PR19 models as set out earlier in this section.

Table 6 Priority areas of modelling agreed with clients and

Aspect of modelling	Issues prioritised	Matters outside scope of study or not prioritised
Timeframe of data used	To use the full historical dataset, which is what we expect Ofwat to use in the absence of strong evidence otherwise. This covers annual data for the period 2013/14 to 2021/22.	To not develop a detailed review of how modelling results vary according to the choice of historical period used for the input data.
Level of aggregation	To consider both aggregated models of total residential retail costs and disaggregated models for bad debt related costs and other retail costs taken separately. To assess the evidence as to whether one type performs better overall than the other. In developing models with alternative drivers, to give more attention to disaggregated models as these tend to give greater chance of cost drivers working in terms of estimation results.	To not prioritise consideration of other ways to disaggregate costs across models (e.g. separate models for meter reading costs).

Aspect of modelling	Issues prioritised	Matters outside scope of study or not prioritised
Aggregate cost or unit cost for the dependent variable	To focus on unit cost models.	Given Ofwat's approach at PR19, and wider considerations, to not consider models in which the dependent variable is an aggregate rather than unit cost measure.
Choice of denominator in unit cost models	To give attention to different options for the denominator, exploring models of cost per customer supplied (cost per household) as well as models of cost per service (where supplying water and wastewater to the same customer counts as two units of service). For models of bad debt related costs, to consider models where the dependent variable is expressed in terms of the ratio of bad debt costs to billed revenue, though we see this as less of a priority.	To not prioritise consideration of using billed properties rather than connected properties, which Ofwat suggested in a cost assessment working group in August 2021. We would expect a high correlation between billed properties and connected properties so would not expect major impacts on the modelling results.
Cost drivers	To consider each of the explanatory variables used for cost drivers in the PR19 models, looking in particular at their robustness in the updated dataset. In prioritised cases (e.g. for deprivation/areas risk variables), to consider whether modelling results can be improved by different functional forms or ways to incorporate cost driver data into a model specification. To explore models that include an explanatory variable that could capture potential differences in reported costs between retailing water and wastewater services (this is something that we found useful as part of work earlier in the PR24 process).	For deprivation/areas risk variables, to focus on the updated data published b Ofwat in November 2022 – including Equifax-based variables that Ofwat did not draw on at PR19 – and not to source alternative data at this stage (e.g. alternative ONS data on income deprivation scores, a wider set of data from Equifax or data from other credit reporting agencies). For the variables that Ofwat constructed from geographically-granular data, to focus on use of the company-level aggregates reported by Ofwat in its recently reported dataset. For the transience/migration variable used at PR19, to use the updated data from Ofwat or updated data from the ONS and to not prioritise looking for alternative data sources for customer transience. To not prioritise trying to include variables to capture any of the following cost drivers in the retail model specifications • Quality of service provided (wholesale and retail) • Relative ease of taking meter readings at customer premises • Regional wages and input prices • Digital literacy and usage • Proportion of off-grid wastewater (or water) customers
Modelling changes over time	To give attention to the potential use of year-specific dummy variables and of time trends in the model specifications, including the types of evidence that can help assess the merits of these approaches relative to the PR19 approach of a constant term only.	To not prioritise considering the case for potentially more sophisticated approaches to capturing dynamic factors which we see as less of a priority at this stage.

Aspect of modelling	Issues prioritised	Matters outside scope of study or not prioritised
Pooled OLS vs random effects	To focus on random effects models, consistent with Ofwat's approach at PR19 and its preference for PR24.	
Smoothing of depreciation over time	To carry out the bulk of our analysis using data that have not been subject to smoothing of depreciation data which Ofwat used at PR19. To cross-check with an approach involving smoothed data to help gauge sensitivity to this.	
Allocation of depreciation and recharges	While not a priority, to check whether the modelling results for the other retail costs models can be improved by using an approximate allocation of some depreciation and recharges to bad debt cost.	

4: Models of bad debt related costs

This section sets out results and analysis on models of bad debt related costs.

We have structured this section so that we present up-front the results of the suite of models we are putting forward. Those models follow from the analyses we carried out over the course of this study, which, in turn, reflects the set of priorities agreed with the client companies (as outlined towards the end of section 3) and our approach to model development and assessment (as outlined in section 2).

The remaining parts of this section are concerned with setting out the analysis underpinning our selection of those models. They consider in turn:

- The choice of the dependent variable.
- Cost driver variables relating to deprivation and arrears risk.
- Cost driver variables relating to the average bill.
- Cost driver variables relating to customer transience.
- The specification of the model dynamics.

Model results

We set out below the 12 models of bad debt related costs that we included in the suite of models produced for this project.

We report these models over two tables, Table 7 and Table 8. The first of these tables present the six models (models labelled D1 through to D6) where the dependent variable is expressed in terms of cost per household; whilst the second table reports on the six models (labelled D7 through to D12) where the dependent variable is expressed in terms of the ratio of bad debt related costs to revenue. All models are panel data random effects models.

In each of the two tables, we present the estimated coefficient for the relevant explanatory variables included in the model and, below each of these and within brackets, the value of the t-statistic. The table also shows, in the row labelled "Dynamics", the dynamic specification assumed for the model: "Dummy vars." Indicates that the model included a set of year-specific dummy variables, and "Trend plus 2020-2022 dummies" indicates that the model included a time trend plus a set of dummy variables for each of the three years in the period from 2019/20 to 2021/22. We have not included in the table the estimated coefficients on the year-specific dummy variables.

We follow Ofwat's convention of using asterisks to flag the statistical significance of the estimated coefficients: * indicates significance at the 10% level; ** at the 5% level and *** significance at the 1% level.

Table 7 Models of bad debt related costs (1)

Model Ref	D1	D2	D3	D4	D5	D6
Dependent variable	Ln(Bad debt related costs per household, CPIH adjusted)					
Explanatory variables						
Ln (revenue per household)	0.971*** (9.569)	0.999*** (10.158)	0.953*** (9.645)	0.975*** (10.564)	1.000*** (11.050)	0.959*** (10.907)
Credit risk score (eq_rgc102)		-0.031 (-1.590)			-0.029** (-2.033)	
Income deprivation score (unadjusted)	0.046* (1.671)			0.046* (1.700)		
Squared income deprivation score			13.395* (1.871)			13.286* (1.880)
Time trend				-0.045*** (-2.701)	-0.047*** (-2.733)	-0.046*** (-2.744)
Dynamics	Dummy vars.	Dummy vars.	Dummy vars.	Trend plus 2020–2022 dummies	Trend plus 2020-2022 dummies	Trend plus 2020–2022 dummies
Overall R-squared	0.74	0.73	0.73	0.74	0.73	0.72
Observations	153	153	153	153	153	153

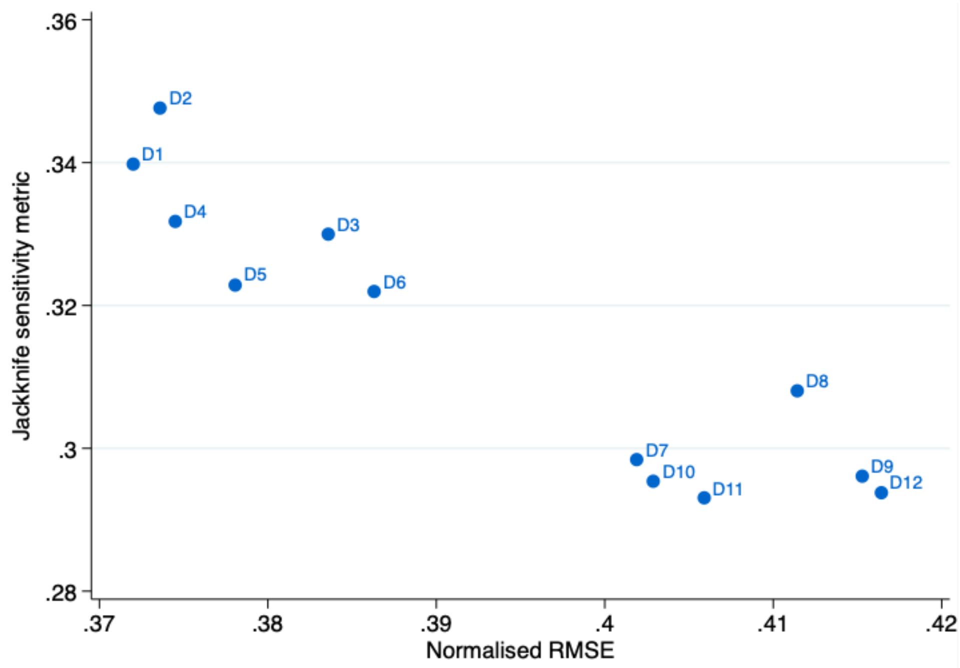
Table 8 Models of bad debt related costs (2)

Model Ref	D7	D8	D9	D10	D11	D12
Dependent variable	Ratio of bad debt related costs to total billed revenue					
Explanatory variables						
Credit risk score (eq_rgc102)		-0.001 (-1.427)			-0.001** (-2.010)	
Income deprivation score (unadjusted)	0.002* (1.887)			0.002* (1.916)		

Squared income deprivation score			0.434* (1.755)			0.434* (1.778)
Time trend				-0.002*** (-3.378)	-0.002*** (-3.403)	-0.002*** (-3.394)
Dynamics	Dummy vars.	Dummy vars.	Dummy vars.	Trend plus 2020–2022 dummies	Trend plus 2020-2022 dummies	Trend plus 2020–2022 dummies
Overall R-squared	0.24	0.21	0.21	0.23	0.21	0.21
Observations	153	153	153	153	153	153

We discussed briefly in section 2, and discuss in greater detail in Appendix 1, the metric on sensitivity of modelled unit costs and the metric on normalised RMSE relating to goodness of fit. Figure 7 shows a chart of the sensitivity against the normalised RMSE metric for the set of models set out above in the two tables above.

Figure 7 Jackknife sensitivity of modelled costs and normalised RMSE: bad debt related cost models



The two clusters of models shown in Figure 7 relate to those where the dependent variable is expressed in terms of cost per household, models D1 through to D7, and to those where the dependent variable is the ratio of cost to revenue, models D8 through to D12. As shown, the first group of models have a lower normalised RMSE, and so a better goodness of fit, than the second group of models. However, the modelled unit costs from those models are more sensitive to variations in the data than is the case for those that model the ratio of costs to revenue, as shown in the figure.

As part of our assessment, we considered the set of tests and analysis covered by the model template submission which Ofwat published in November 2022. In this regard, we note that the set of models D1 through to D6, modelling cost on a cost per household basis, fail the RESET specification test. This result is indicative of potential opportunities for a model variant (e.g. involving quadratic or interaction terms) with a more complex functional form that could fit the data better. That said, in practice there may be limited opportunities to explore variations with such added complexity, given the small sample size, and it is possible that such variants would perform less well on other dimensions which are of interest, namely on the sensitivity around modelled costs.

Dependent variable

We have considered models of bad debt related costs with two alternative specifications of the dependent variable:

- bad debt related costs per household (in logarithmic terms); and
- ratio of bad debt related costs to household retail revenue.

We think there are good grounds to consider each of these.

Ofwat used the first of these specifications in its bad debt related costs models at PR19.

The motivation for the second specification is that household retail revenue may be regarded as an appropriate scale variable with respect to bad debt costs and so be an appropriate metric to use for normalising bad debt related costs when comparing companies. We expect there is operational support for this alternative formulation, namely:

- We expect that the level of provision of bad debt that is made by companies, which represents the largest components of bad debt related costs, to be related with the amount of retail revenue billed, and so at risk, for a given profile of customer risk.
- In respect of debt management costs, the other component of bad debt related costs, we would also expect that the costs companies choose to incur to manage debt, namely to pursue unpaid debt, to also be driven by the scale of the revenue at risk.

Figure 8 and Figure 9 show the variation across companies for each of the two normalisations of bad debt related costs. The values presented are the average over the five years from 2017/18 to 2021/22. In each of the figures, the average across all companies is shown by the grey horizontal line.

Figure 8 Bad debt related costs per household (£/customer, CPIH adjusted, 2017/8 to 2021/22)

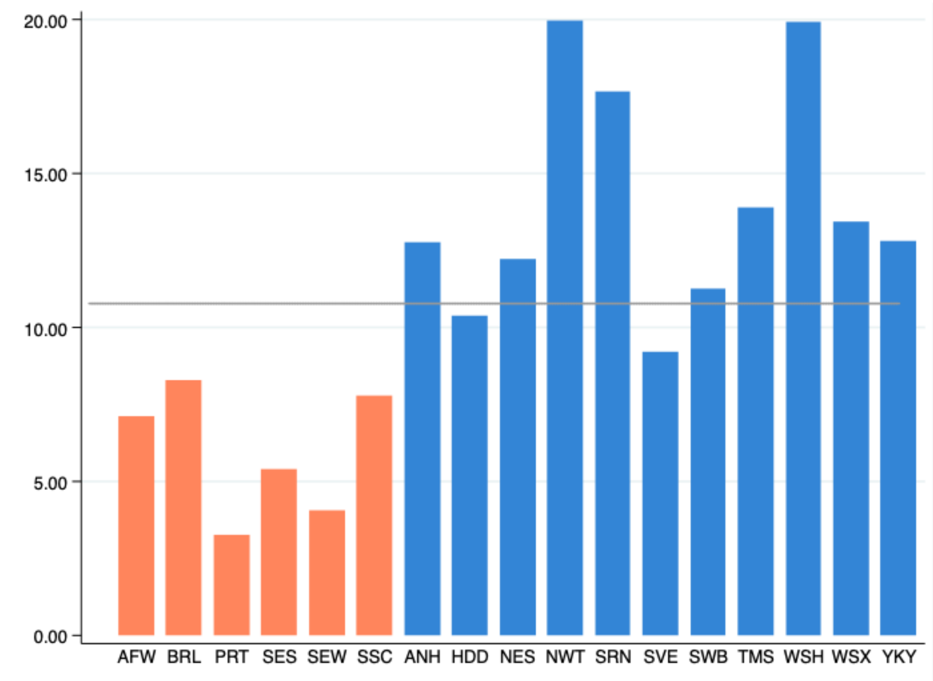
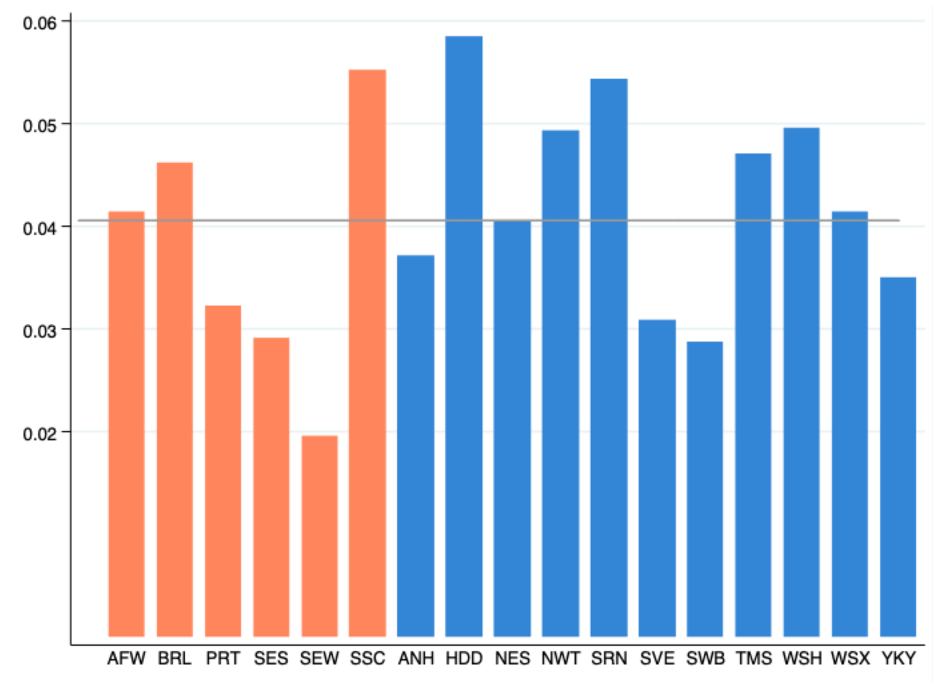


Figure 9 Ratio of bad debt related costs to household retail revenue (2017/8 to 2021/22)



Contrasting the two charts above, it seems clear that, when comparing costs on a per household basis (shown in Figure 8), there are material differences between water only and water and wastewater companies. In the period since 2017/18, the average bad debt related cost per household was £14.0 for water and wastewater companies, and £7.0 per household for water only companies. A candidate factor that helps explain that difference is the variation in the size of the average household bill across companies. Ofwat's models at PR19 reflected this and incorporated an explanatory variable capturing this in its models of bad debt costs. As discussed in the earlier section reviewing the PR19 models that variable was found to have an important role in explaining the observed differences across companies' bad debt costs per household.

In the light of our quantitative analysis, we found a good case for including models in which the dependent variable is specified as the ratio of bad debt related costs to total billed revenue, alongside models of bad debt cost per household (as used for PR19). The former is a more restrictive model specification, imposing an assumption of a 1:1 relationship between a company's bad debt costs and the average size of its bills to customers, but this seems quite a reasonable assumption from an economic perspective, at least as an approximation. Given the small sample size and complexities of companies' cost structures, there can be benefits from imposing such an assumption rather than allowing the relationship to be estimated by the model. As shown in Figure 7 earlier in this section, our analysis indicates that the results for models involving this assumption were considerably less sensitive to variations in the dataset than corresponding models where the dependent variable was expressed on a cost per household basis, though the goodness of fit, as measured by the normalised RMSE metric, was slightly lower. In terms of the range of efficiency ratios derived from the models, and which may also be regarded as an alternative metric related to goodness of fit – though more rudimentary than the normalised RMSE metrics – we note that models D7 through to D12, where the dependent variable is the ratio of cost to revenue, tended to produce a slightly narrower range than models D1 to D6, where the dependent variable is based on the cost per household.

Deprivation and arrears risk

The set of models for bad debt related costs and for total retail costs which Ofwat drew on at PR19 included an explanatory variable intended to reflect factors that drive households to fall behind payment of their water bills. The models did so through the inclusion of two different metrics within the set of explanatory variables, namely:

- A metric capturing the average income deprivation score across the geographic area served by each water company. This was derived from data from the ONS and by Statistics for Wales.
- A metric capturing the proportion of households within the geographic area served by each water company which had a default, drawing on data from the credit reference agency Equifax.

There are sound intuitive reasons for controlling for variations in the propensity of households to fall behind with their water bills, or to not pay them at all, in models of bad debt related costs. Our analysis supports this, echoing Ofwat's findings at PR19.

In our work, we explored a set of alternative metrics to capture deprivation or arrears risk. These drew mostly on the dataset published by Ofwat in November 2022, as set out in Table 9.

Table 9 Metrics on deprivation or arrears risk explored

Metric	Source
Equifax – Insight Postcode Event: Percentage of households with default (eq_lpcf62)	Ofwat November 2022

Metric	Source
Equifax – Credit risk score derived from all Insight data (eq_rgc102)	Ofwat November 2022
Equifax – Avg. number of Partial Insight accounts or county court judgements (eq_xpcf2)	Ofwat November 2022
Combined Income Score for England and Wales (IMD) - unadjusted	Ofwat November 2022
Council tax collection rate	Ofwat November 2022
Weighted average of the square of income deprivation score	ONS and Statistics for Wales

The last metric listed in the table – labelled as the weighted average of the square of income deprivation score – is one which, unlike the others, we derived ourselves, drawing on data at the LSOA-level published by the ONS and Statistics for Wales. We describe in Appendix 2 details about how we constructed that variable. In broad terms, the measure is a weighted average of the square of the income deprivation score across the LSOAs that fall within each company's service area. We considered the square term of the local-level income deprivation score with a view to exploring the hypothesis that there is a non-linear relation between the measure of income deprivation and the propensity to fall behind, or not pay, water bills.

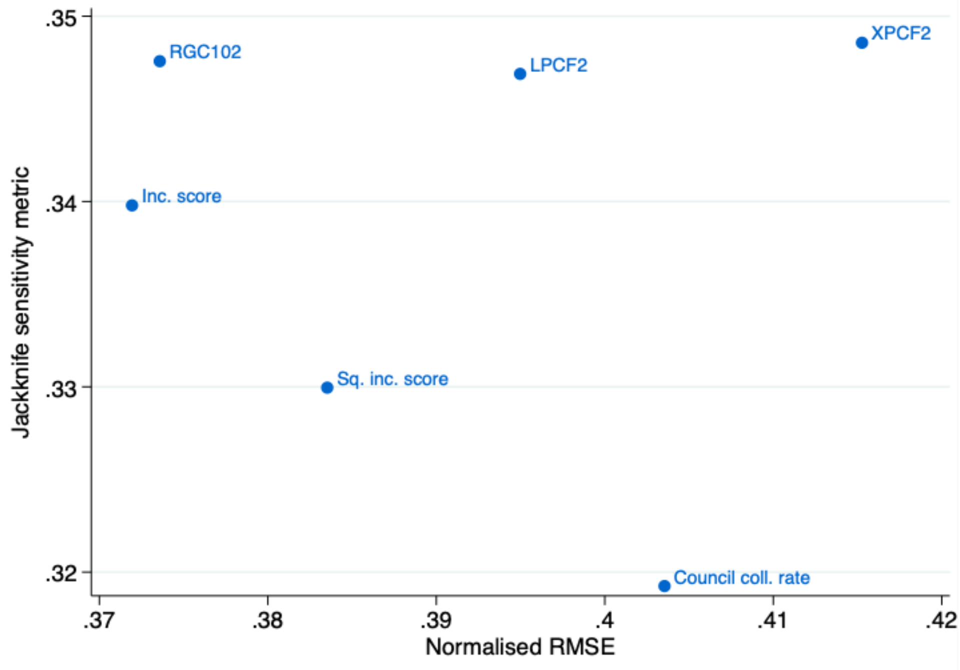
The rationale for the hypothesis may best be seen through a simplistic example. Consider two companies, each serving just two LSOAs. The income deprivation score in the two LSOAs served by company A are 20% and 30%, whilst the score in each of the LSOAs served by company B is 25%. Consider the case where the geographically-granular data are aggregated to the company level in a 'linear' way, the approach followed, to the best of our understanding, by Ofwat in producing the company-level measures of the income deprivation score and of the different Equifax-based variables. Under such an approach, the company-level metric for the two companies would be the same, 25%. On the other hand, squaring the LSOA-level deprivation scores ahead of aggregating them to company level gives, in effect, greater weight to LSOAs with higher deprivation scores. Under such an approach, the company-level metric for company A would be greater than that for company B.

We applied the approach of aggregating the square of LSOA-level income deprivation data to the measures published by the ONS as this is the measure for which we had geographically-granular data. The data from Equifax which were published by Ofwat in November 2022 were already at a company-level and more granular data were not set out so we could not adopt the approach we took in relation to the income deprivation score data.

We explored the role of each of the measures set out in Table 9 in explaining variations across companies' bad debt related costs. We found that both of these metrics were statistically significant (at, or below the 10% significance level), and their estimated coefficients were of the expected sign, i.e. companies serving household with greater income deprivation score or higher measures of arrears risk are predicted, by the models, to have greater bad debt related costs. On the other hand, when we controlled for the variation in arrears risk through the inclusion of the Equifax measure relating to proportion of default (LPCF2) or through number of Partial Insight accounts of county court judgements (XPCF2) we found those variables to not be statistically significant.

We compared the performance of models including the alternative measures of deprivation or arrears risk in respect of their goodness of fit (measured by the normalised RMSE metric) and the sensitivity of the modelled unit costs, following the approach we outlined in section 2 and set out in more detail in Appendix 1. Figure 10 presents a chart of such analysis for a set of models of bad debt costs that are identical except for the choice of metric used to control for the variation of deprivation or arrears risk across companies. For ease of interpretation, we have used that variable name to label the position of each of the models.

Figure 10 Jackknife sensitivity of modelled costs and normalised RMSE of alternative models of bad debt related costs



Taking the model that includes the Equifax variable RGC102 as a reference point, the figure shows that the models that include either of the other two Equifax variables – XPCF2 or LPCF2 – perform less well in terms of goodness of fit, with no or very marginal improvement with regard to how sensitive the modelled costs are to variations in the dataset In contrast, the model including the ONS income score deprivation performs marginally better in respect of goodness of fit, as measured by the normalised RMSE, – consistent with the marginally higher R-squared statistic for model D1 compared to that of D2 (0.74 versus 0.73) shown in Table 7 – and performs a bit better in respect of the sensitivity of modelled costs. For the models drawing on the square of the income deprivation score and on the council tax collection rate, the figure suggests a trade-off between goodness of fit and sensitivity.

Going beyond considerations of its statistical performance when included in models, we have a concern with regards to the use of data on council tax collection rate as a means of proxying the variation in deprivation or arrears risk. In particular, we would expect that differences in council tax collection rates reflect not just the variation in households' ability or attitude to paying their council tax bills but also other factors, such as the differences in the effectiveness and approaches of councils managing and pursuing unpaid council tax. This is a concern if the extent of such differences was not independent of the underlying factors of interest to us, namely ones related to deprivation or risk of non-payment of bills.

Reflecting the areas to prioritise in the study, presented earlier in Table 6 in section 3, we did not pursue other approaches to controlling for deprivation or arrears risk, such as exploring the development and use of a composite variable combining several of the measures we considered separately or exploring metrics drawing on other data altogether.

Of the three Equifax variables included in the retail dataset published by Ofwat in November 2022, we found that the variable RGC102, a measure of average risk arrears risk performed better statistically in models – in terms of the size of its t-statistic and in terms of the goodness of fit and sensitivity of modelled costs in models that included that variable – than the other two Equifax measures. We also consider that, intuitively, it is a more appealing one to draw on for the modelling, as the metric reflects draws on a richer set of information, processed by Equifax, than either of the other two measures. The RCG102 measure is a different one to that used by Ofwat at PR19 but is the one that we had given emphasis to in work we carried out for United Utilities at PR19.[7]

We also found a good case for models where deprivation is controlled for through the inclusion of the ONS measure on income score deprivation, both when this is aggregated linearly – which is the approach taken in the measure published in Ofwat's retail dataset – and when it is aggregated across geographically-granular units (LSOAs in the case of the data we used) after first taking the square of the measure. In either of these formulations, we found the income score derivation variable to be consistently significant and that models which included either of them to perform well, relative to the other measures we considered, in respect of goodness of fit and sensitivity of modelled costs.

Average bill

Our analysis supports the inclusion of the average household bill within the set of explanatory variables in models where the dependent variable is defined in terms of bad debt costs per customer. This is in line with the models put forward by Ofwat at PR19.

In the models we explored, the estimated coefficient on that variable is broadly around 1. Such an estimated coefficient indicates that, say, a 10% difference in the average bill between two companies is associated with a 10% difference in bad debt cost per household. We find that the t-

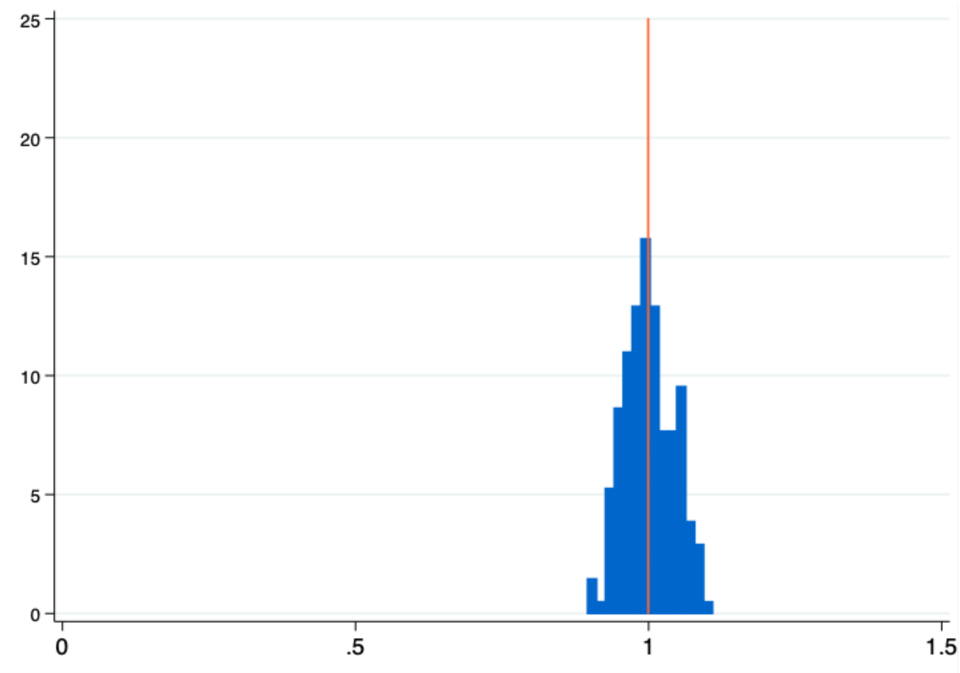
---

[7] See Reckon (2017) "A working
to explore how socio-economic deprivation and arrears risk can be captured in models of residential retail cost", accessible from https://www.unitedutilities.com/globalassets/z_corporate-site/about-us-pdfs/looking-to-the-future/depriviation-and-arrears-risk-in-hh-retail-cost-assessment-100517.pdf.

statistic on the estimated coefficient for this variable is systematically high, across different models we looked at, pointing to the variable being significant with a very high confidence level. Figure 11 shows the histogram for the estimated coefficient of this variable for one of the models when that model is estimated using different variations of the dataset, an approach we outlined in section 2 and which we explain more fully in Appendix 1. As shown, the estimated coefficient on the variable related to average bill is consistently found to be within a narrow range and very close to one, across the set of dataset variations. The figure is illustrative of the high degree of robustness of the estimated coefficient for this variable to variations in the dataset.

In light of our analysis, we considered that there was a very strong case for including the average bill variable in all models within our suite for which the dependent variable was defined on the basis of bad debt related costs per household.

Figure 11 Histogram of estimated coefficient for Ln(Average bill) in model D2



Customer transience

We explored the role of controlling for population transiency, drawing on the variable included in Ofwat's Autumn 2022 dataset.

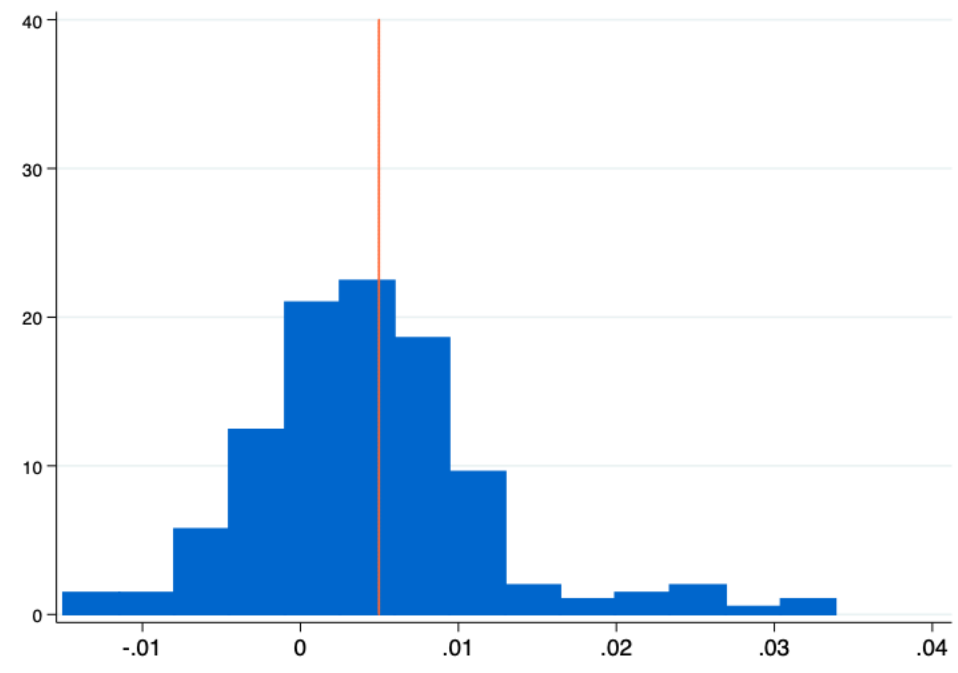
The rationale for considering such a variable is that, plausibly, population transiency is associated with bad debt related costs as it may be harder for companies to pursue unpaid bills of customers that have moved home. In its PR19 models, Ofwat included this variable in one of its models of bad debt related costs.

The metric that Ofwat uses to capture population transiency is based on data on migration into or out of Local Authority Districts (LAD). As such, the metric does not capture household moves within an LAD, which could account for a significant share of population flows. However, we are not aware of publicly available data that would capture such movements within an LAD and, we too, drew on the metric that Ofwat produced and published.

We explored the role of that metric by including it in a set of models of bad debt related costs that we considered. We examined the performance of such models compared to ones that did not include the transiency variable. The results are slightly mixed but, by and large, did not provide strong grounds for the inclusion of the variable in models:

- The sign on the estimated coefficient for the transiency variable was different depending on the choice of metric used to capture income deprivation / arrears risk. For example, in models that included the ONS measure of income deprivation score, the estimated coefficient on the variable was positive – in line with economic/operational intuition – but in models that included the Equifax LPCF2 or RGC102 variable it was negative, which is counter-intuitive.
- Across the set of models explored, we found that the estimated coefficient on the transiency metric varied across a relatively wide range when a given model was subject to the jackknife analysis. Further, that range straddled both positive and negative values, particularly where the model included the LPCF2 or RGC102 Equifax variable. This is shown in Figure 12, a histogram of the estimated coefficient of the variable proxying transiency when this is included in a model which, save for this, has the same specification as model D2.
- The inclusion of the transiency measure (necessarily) improved the R-squared measure of goodness of fit, though the incremental impact was small, and tended to increase significantly the variance of the predicted unit costs.
- The above findings held for models where the dependent variable is expressed on a cost per household basis, and when it is expressed as the ratio of bad debt related costs to revenue.

Figure 12 Histogram of estimated coefficient for measure of transiency in variant of model D2

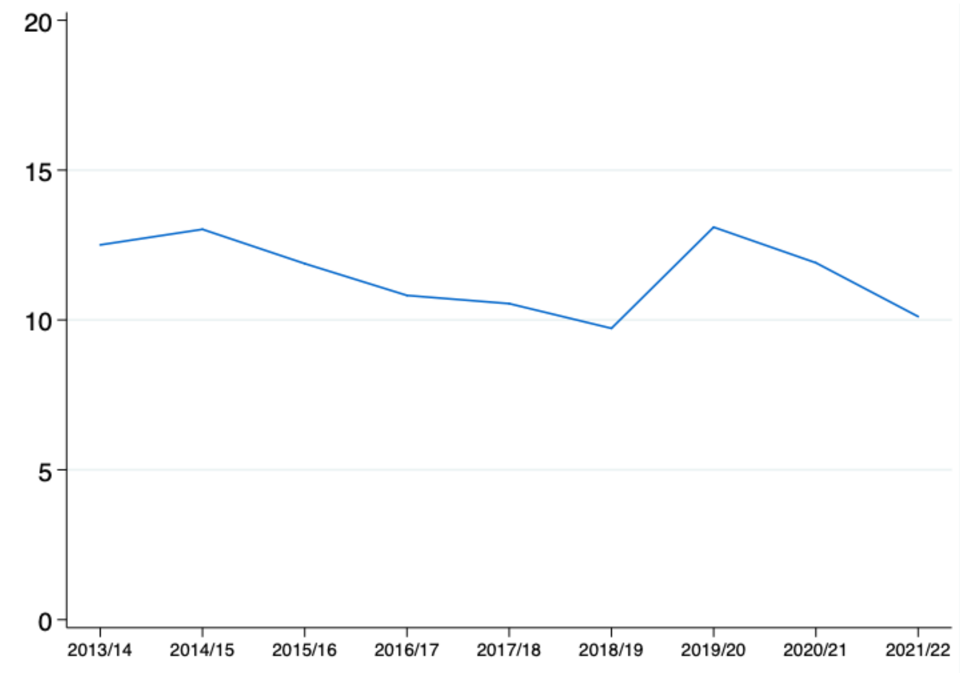


Drawing on the results of our analysis, we did not see good grounds to include a variable capturing customer transiency, as proxied by the metric derived from migration data, in any of the models within the suite of models to put forward. The estimated coefficient on the variable tended to have low t-statistics and in some of the model variants, the sign was negative, which is counter-intuitive. Its inclusion in models made a marginal contribution to improving the goodness of fit but increased the sensitivity of modelled costs.

Model dynamics

An area we prioritised in the development of models concerned the specification of the dynamics in the models. One of the motivations for this was the observation that industry's costs have varied in the period 2013/14 to 2021/22, the period spanned by the dataset which Ofwat proposes to draw on. This is illustrated in Figure 13 which charts the average of water companies' bad debt related costs per household, in real terms.

Figure 13 Industry-average bad debt related cost per household, 2013/14 to 2021/22 (£/ household, 2017/18 CPIH adjusted)



For the above graph, and throughout our analysis, we have used the measure of bad debt related costs which takes account of the ex-post adjustment made by water companies in respect of the bad debt loss provisions they had made in the earlier phase of the Covid-19 pandemic.

Figure 13 suggests a downward trend in real unit bad debt related costs in the period up to 2018/19, followed by a significant increase in 2019/20, in the early phases of Covid-19 period, and a reversion towards the pre-Covid-19 trend for 2021/22.

Motivated in part by the observed trend in the chart, we considered alternative specifications for model dynamics, namely:

- For purposes of comparison, models that kept the static specification of Ofwat's PR19 models.
- Model specifications with a time-trend explanatory variable.
- Model specifications that included a set of year-specific dummy-variables, e.g. a dummy variable that is equal to 1 if an observation refers to 2013/14 and is 0 otherwise, and similarly for all, but one, of the other years in the sample period.
- Model specifications with a time-trend and a set of year-specific dummy variables for each of the years from 2019/20 onwards.

The last of these specifications was aimed at exploring whether such a structure was necessary to control for the apparent change after 2018/19 in the trend of unit costs, at an industry-average level.

We found strong support for the last two specifications for the dynamics in models.

Compared to the last two specifications on the dynamics outlined above, we found that models allowing no dynamics or imposing a constant time trend throughout the period showed a higher prevalence of variables with estimated coefficients that were not significant and tended to perform less well with regard to goodness of fit and sensitivity of modelled unit costs.

For the models where a time trend was included, models D4 to D6 and models D10 to D12 in Table 7 and Table 8 we found that the estimated value for its coefficient was – 0.04 in the case of models expressed on a cost per household basis, and around –0.002 in the case of models where the dependent variable was defined as the ratio of costs to billed revenue. The results for the first set of models indicate a year-on-year decrease in bad debt related costs per household of around 4% in real terms, having controlled for cost driver explanatory variables; the results for the second set indicate a year-on-year fall in the ratio of costs to revenue of 0.002 points, which is around 5% of the industry average ratio (over the sample period).

The estimates just discussed are of the trend across the period spanned by the data, 2013/14 to 2021/22. The models that had a time trend in their specification also included a set of year-specific dummy variables, one for each of the period since 2019/20. We found that dummy-variables for 2019/20 and 2020/21 were consistently positive and statistically very significant, and that this was not always the case for the dummy variable controlling for 2021/22 effects.

For the set of models we put forward in our model suite, we include ones where the dynamic aspects are captured through (i) the inclusion of a set of year-specific dummy variables for each year in the sample, and through (ii) the inclusion of a time trend plus a set of dummy variables for each of the last three years in the sample. We found these two specifications of the dynamics to perform better, in terms of the goodness of fit od models, and sensitivity of modelled costs, than the alternatives we looked at, including the alternative of not modelling any dynamics at all (as is the case of the PR19 models). With regard to the specification where we include a time-trend plus a set of dummy variables for the three years since the start of the Covid-19 pandemic, further consideration may be given to the case for including a dummy variable in respect of 2021/22, the most recent of those three years. We have kept that in for the suite of models developed here, but this could be something to explore further, in particular once data for 2022/23 become available.

5: Models of other retail costs

This section sets out results and analysis on models of other retail costs.

The section follows an analogous structure to the previous one. We present up-front the results of the suite of models of other retail costs we are putting forward. These follow from the analyses we carried out over the course of this study and which, in turn, reflect the set of priorities we agreed with the client companies (as outlined towards the end of section 3) and our approach to model development and assessment (as outlined in section 2).

The remaining parts of this section are concerned with setting out the analysis underpinning our selection of those models. This considers in turn:

- The choice of the dependent variable.
- A potential cost driver explanatory variable relating to the proportion of dual service customers.
- A potential cost driver explanatory variable relating to the mix between water and wastewater retail services.
- Potential cost driver explanatory variables relating to meter penetration.
- Potential cost driver explanatory variables relating to economies of scales.
- The role of each of the cost drivers explored in the analysis.
- The specification of the model dynamics.

Model results

We set out in Table 10 the six models of other retail costs that we included in the suite of models produced for this project. The table is structured in the same way as those in which we presented the models of bad debt related costs in section 4.

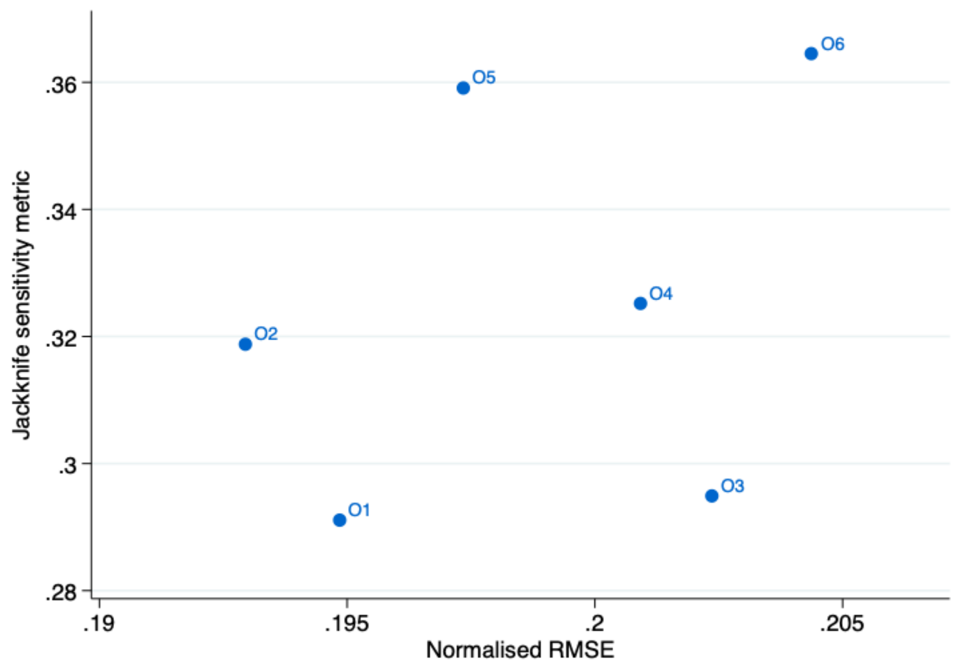
Table 10 Models of other retail costs

Model Ref	O1	O2	O3	O4	O5	O6
Dependent variable	Ln (Other retail costs per household, CPIH adjusted)				Ln (Other retail costs per service, CPIH adjusted)	
Explanatory variables						
Proportion of dual service customers	0.208** (2.367)	0.286*** (3.597)	0.210** (2.395)	0.292*** (3.562)		
Proportion of metered connections	0.495* (1.709)	0.509* (1.798)	0.555** (2.045)	0.577** (2.163)		

Ln (Total households connected)		-0.033 (-1.053)		-0.035 (-1.095)			
Prop. of measured services						0.603* (1.938)	0.663** (2.244)
Prop. of services that are wastewater						-0.717*** (-4.431)	-0.718*** (-4.558)
Time trend			-0.020* (-1.945)	-0.020* (-1.951)			-0.021** (-2.093)
Dynamics	Dummy vars.	Dummy vars.	Trend	Trend	Dummy vars.	Trend	
Overall R-squared	0.18	0.20	0.15	0.16	0.48	0.46	
Observations	153	153	153	153	153	153	

Figure 14 charts the performance of the set of models reported above in terms of the jackknife sensitivity of modelled costs and in terms of their goodness of fit, as measured by the normalised RMSE metric.

Figure 14 Jackknife sensitivity of modelled costs and normalised RMSE: other retail cost models



The scatter chart in Figure 14 shows that the models that include a set of year-specific time dummies, models O1, O2 and O5, have, as would be expected, a better goodness of fit. The modelled unit costs associated with those models are also less sensitive than the respective model variants (models O3, O4 and O6) when, instead, the model specification included a time trend to capture changes over time. The figure also points to the finding that for the two models where the dependent variable is expressed in terms of cost per service, models O5 and O6, modelled costs are particularly more sensitive than is the case for models that benchmark cost per household.

We considered the set of tests and analysis covered by the model template submission which Ofwat published in November 2022. In this context, one point to comment is that models O5 and O6 fail the RESET specification test, indicative of potential opportunities for a model variant (e.g. involving quadratic or interaction terms) with a more complex functional form. In practice, however, there may be limited opportunities to explore models with such added complexity, given the small sample size, and it is possible that such variants would compromise the model's performance on other dimensions of interest, namely on the sensitivity around modelled costs.

Dependent variable

We explored two alternative specifications for the dependent variable:

- other retail costs per household; and
- other retail costs per service.

At PR19, in its models for other retail costs, Ofwat expressed the dependent variable as the natural logarithm of the cost per household, the first specification listed above. We think there are also grounds to consider models where the dependent variable is expressed in terms of the cost per service provided. In contrast to the use of households as the variable to normalise costs, using the number of services provided gives, in effect, twice the weight to households to whom a company provides both water and wastewater services rather than just one of those services.

The rationale for this alternative specification stems from our understanding of the type of arrangements that exist between some water and wastewater companies in respect of retail services. In particular, we understand that, though not in place across the whole industry, there are generally arrangements such that a water-only company will provide customer services, meter reading and potentially debt management services on behalf of the company providing wastewater services in their area of service. Where such an arrangement is in place, the cost figures reported by the water-only company will, as required by the RAGs, exclude the costs incurred in providing such services to (or on behalf of) the wastewater company; those will be reported by the wastewater company in their own cost tables. Where the same company provides both water and wastewater services, then the full costs associated with providing those two services will be reported within that company's cost figures. In the case of Bristol Water and Wessex Water, the Pelican joint venture, which carries out retail services on behalf, gives rise to a similar effect.

The existence of such arrangements lends support, we suggest, to considering that from an economic and operational point of view, the number of services provided may be a more appropriate scale variable for other retail costs than the number of customers.

Expressing the dependent variable on a cost per service basis rather than on a per household basis has an impact on the candidate set of explanatory variables to consider, most significantly on the need to control for the proportion of dual service customers. We discuss this, and the role of other candidate explanatory variables below.

Whilst, as discussed above, we consider there are good intuitive grounds for considering models where the dependent variable is defined on the basis of cost per service, rather than on the basis of cost per household, we were surprised to find that the 'per service models' did not perform as well in terms of goodness of fit, as measured by the normalised root mean square error metric presented in section 2, or in terms of the sensitivity of the predicted unt costs. We note, however, that, as was set out earlier in Figure 4, the spread of the efficiency ratios – which also captures a measure of the goodness of fit – for the models expressed on a cost per service basis are not wider than those derived for models of cost on a per household basis.

On balance, we chose to include within our suite of models ones where the dependent variable is defined in terms of cost per service.

In terms of their statistical performance, there is a case for preferring models of cost per household than ones of cost per service. That said, we considered there were a number of other relevant factors in favour of models expressed on a cost per service basis. First, a cost per service specification allowed for models that did not impose the restriction that retailing to a water-only and to wastewater-only customer has the same cost; we did not find a satisfactory specification of a model where the dependent variable was cost per household that allowed for this and which produced robust results (we discuss this further below). Second, we have some concerns about cost per household models in terms of the magnitude of the estimated coefficient on the proportion of dual service household not being consistent with economic / business expectations. Neither modelling approach is perfect in this regard – capturing relative cost of single versus dual service households – but this issue seems more of a concern in models of cost per household compared to models of cost per service. Third, we consider there is some benefit from having some diversity in the suite of models.

Proportion of dual service customers

At PR19, Ofwat's models of other retail costs, as well as its models of total retail costs, included the proportion of dual service customers within the set of explanatory variables. The dependent variable in those models was the natural logarithm of the cost per household. Ofwat provided the following reasoning for controlling for the proportion of dual service customers:[8]

> *"Dual service customers receive both water and wastewater services from the same company. Dual customers may generate more contact and enquiries relative to single service customers, which in turn drives customer service costs. We expect such incremental costs associated with dual customers to be small.*
>
> *We include the proportion of dual customers in our other retail costs models. The variable is statistically significant, positive, and small in magnitude as expected."*

---

[8] Ofwat (2019) "Supplementary technical appendix: Econometric approach".

Whilst we agree that there are very good economic reasons for including a variable capturing the proportion of dual service customers in those models, these do not stem from the observation made by Ofwat as captured in the quote above.

Ofwat's reasoning is expressed in terms of the incremental costs associated with providing retail services to dual service customers, compared to single service customers. In the light of the widespread arrangements that are in place between water and wastewater companies, we think that that is not the most appropriate way to view things.

Instead, we suggest that the basis for controlling for the proportion of dual service customers – in models where the dependent variable is expressed on a cost per household basis – is that for dual service customers the water retailer supplying them does not have the ability to share with another company part of the retail costs associated with serving them.

To illustrate this, consider the following example. Assume that the cost of providing water and wastewater retail services is £20 per customer. Water-only company *A* carries out retail services on its own behalf and on behalf of the wastewater company *B,* who is the wastewater provider in *A*'s area of service. Company *A* charges company *B* for this – let us say it charges £10 per customer. On the other hand, in another region, company *C* provides water and wastewater services to all its customers and the £20 per customer of retail costs that it incurs are not shared with any other company. Under these assumptions:

- A comparison of those companies' costs per service would show that companies *A, B* and *C* each have a cost per service supplied of £10.
- A comparison of those companies' costs per customer would show company *A* having a cost per customer of £10, whilst company *C* would have a cost per customer of £20.

This simplified example makes it clear why it would be necessary for a comparison of companies' costs on a per customer basis to take account of the proportion of dual service customers. The reason is not concerned with potential incremental costs associated with serving dual service customers (e.g. more contact and enquiries relative to single service customers as suggested by Ofwat). In our simplified example there is no such issue as in both regions the company providing the retail service is doing so in respect of both water and wastewater services. Rather, the reason for controlling for the proportion of dual service customers is that such variable reflects the extent to which a company can share the costs of serving its retail customer base with another company.

From the analysis, we find that in models of other retail where the dependent variable is expressed on the basis of cost per household, the variable capturing the proportion of dual service customers consistently plays an important role in explaining the observed variation across companies' unit costs.[9]

[9] One issue to bear in mind when interpreting models that include the proportion of dual service customers as an explanatory variable is that that variable is associated with other potential cost drivers, in particular, with scale and mix of services provided. These associations may dampen the ability of a given model to isolate the role of the dual service variable in explaining cost variations across companies.

Overall, we saw a strong case for including an explanatory variable for the proportion of dual service customers in models in which the dependent variable in defined in terms of cost per household. (Where, in contrast, the dependent variable is expressed on a cost per service basis, we do not see that it would be necessary to control for differences in the proportion of dual service customers.)

However, on further analysis, we think that there is some potential concern about the scale of estimated coefficients for the dual service explanatory variable which are relevant to be aware of. This issue applied to the coefficient of that variable in the models we estimated as part of this project as well as in the Ofwat PR19 models.

The estimated coefficient on the variable capturing the proportion of dual service households is around 0.21, in those models not controlling for economies of scale (models O1 and O3), and around 0.29 in models where economies of are controlled for (models O2 and O4). Given the logic underlying the inclusion of this variable, we would have expected a higher coefficient.

One way to illustrate why we would have expected a higher coefficient for that variable is to consider how hypothetical changes in the geographic boundaries of water companies' retail activities – which will have an impact on proportion of dual service customers – affect the costs that would be predicted by different modelling approaches. Consider, for example, that Bristol Water and Wessex Water were to merge and become a single entity rather than separate companies acting through a joint venture. How would the costs predicted by the models for the two separate companies compare with those that would be predicted for the single entity?

Drawing on the model results set out in Table 10, we calculated, for the cost per household model O3, that the total predicted residential retail costs of the merged entity would be around 20% lower than the sum of the costs predicted for the two companies taken separately. This seems counterintuitive: model O3 does not allow for economies of scale and so it would not be expected that the combined predicted costs for providing retail services to the Bristol Water and Wessex Water customer base would be so different depending on whether these are assumed to be supplied by a single company or by two separate companies. This issue is attributable to the estimated coefficient on the dual service variable in that model: the costs of the merged entity would be the same as for the separate companies if the coefficient was higher, specifically if it were around 0.64 rather than 0.21.

In contrast, under model O6, which is a "cost per service" version of model O3, the results are much less sensitive to this issue. We found in our example that model O6 would imply that the aggregate costs for the merged entity would be only 1.7% less than the total costs estimated if each company were considered separately.

While neither modelling approach seems to give perfect results, there seem to greater concerns about this aspect of the results for those models that are on a cost per household basis than for those models where the dependent variable is defined on the basis of the cost per service.

Our concerns about the scale of the estimated coefficients on the dual service explanatory variable is one of the reasons why the model suite we present in this report includes some models which are expressed on a cost per service basis.

The mix between water and wastewater retail services

As part of our model development work, we explored the case for including an explanatory variable to capture the mix between water and wastewater retail services, which is a potential cost driver for a company's residential retail costs.

An implicit assumption of Ofwat's PR19 residential retail models is that the costs of providing retail services (excluding bad debt related costs) to a water-only customer are the same as the costs of providing retail services to a wastewater-only customer.

For models where the dependent variable is defined in terms of cost per household, we were not successful in developing models which included an explanatory variable relating to the mix between water and wastewater retail services. This is likely to be related to the multi-collinearity between this variable and the dual service variable included in these models. As a consequence, our models O1 to O4 share the feature of Ofwat's PR19 models of not allowing for any differences in retail costs between water and wastewater services.

For models where the dependent variable is defined in terms of cost per service, we found that it was possible to control for the mix between water and wastewater services and that doing so improved a model's performance with regard to goodness of fit and with regarding to reducing the sensitivity of modelled costs. In our analysis, we have done this by including in those models a variable defined as the proportion of services provided which are a wastewater service.

We found that this variable was consistently significant, at a significance level below 1%, in models expressed on a cost per service basis. Across those cost per service models, the estimated coefficient for it is estimated to be around – 0.72. The negative sign indicates that these models predict that companies' cost per service are lower the greater the proportion of services that are wastewater services. Put differently, the models predict that the provision of wastewater retail services has a lower unit cost than the provision of water retail services.

There is a corollary to this in in relation to what is implied in respect of the costs of serving dual service customers versus single service customers. Due to the inclusion of a variable controlling for service mix, in those models expressed on a cost per service basis (e.g. models O5 and O6 in Table 10) there is no assumption imposed on the relative costs of dual service versus single service customers or on the relative costs of water-only compared to wastewater-only customers – these relationships are estimated by the model.

To illustrate this, we compared what the models predicted for three hypothetical companies: one serving wastewater-only customers, another serving water-only companies and the third serving just dual service customers. To isolate the effect related to service mix, we assumed an equal level of meter penetration – a relevant explanatory variable in the models – and applied the thought experiment using the regression results for models O3 and O6. These two models do not control for economies of scale and so results are not affected by that driver. In model O3 the dependent variable is expressed in terms of cost per household, and in model O6 it is expressed in terms of cost per service.

For model O3, the implied unit costs of serving water-only households and of serving waste-only household are, by definition of the assumed model specification, the same. The results of the model also imply that serving a dual service household has estimated costs which are 1.23 times the cost of a single-service household.[10]

For model O6, we find that the implied costs of a serving waste-only households are around half of those of serving water-only households and that serving a dual service household has costs of 1.40 times the costs of a water-only household.[11]

These worked examples help illustrate the following:

- Because they include explanatory variable relating to the mix between water and wastewater retail services, there is no assumption/implication of a dual service customer costing twice as much as a single service customer in those models (O5 and O6) where the dependent variable is expressed in terms of cost per service; and
- There are interesting differences between models O3 and O6 in terms of what they imply about the relative costs of retailing to water and wastewater customers.

The costs of providing wastewater services relative to those of providing water services which are predicted by the cost per service models (models O5 and O6) are lower than might have been expected. One hypothesis, which we are not able to test, is that this may be driven by the relative bargaining power of WASCs vs WOCs when setting up agreements for joint billing arrangements (perhaps historically) rather than only reflecting underling cost drivers.

Meter penetration

There is an operational rationale for considering metrics on meter penetration as a candidate cost driver of other retail costs. Most obviously, we expect that the cost of meter reading, one of the elements within other retail costs, is related to the number of metered customers. In addition to this, it also seems plausible that there are cost differences between serving metered and unmetered customers, e.g. arising from differences between the two customer segments with regard to contacts made by customers to the company to enquire about meter readings or bills.

We think it useful to note that, further to the potential operational rationale for it, Ofwat may see that it is useful, from an incentive point of view, for the econometric models to include a metric related to meter penetration as an explanatory variable, provided the coefficient on this is estimated to be positive. In such circumstances, the inclusion of an explanatory variable related to meter penetration will provide a route through which companies can recover the potentially additional costs associated with serving metered customers.

We found that, consistently across models, the coefficient on the variable on meter penetration was statistically significant at the 10% or lower significant level, and that it had a positive sign, as would

---

[10] This figure is the exponential of 0.210, where 0.210 is the estimated coefficient of the variable dual service customers in model O3.

[11] This figure is the 2 times the exponential of (0.5 * –0.719), where –0.719 is the coefficient of the variable capturing proportion of wastewater services. In this calculation, the coefficient is multiplied by 0.5 to reflect fact that, in our illustrative example, we assume that the dual service company only serves dual service households and so its proportion of wastewater services is 0.5. The multiplication by 2 is to express the predicted difference of cost per service into predicted difference in terms of cost per household.

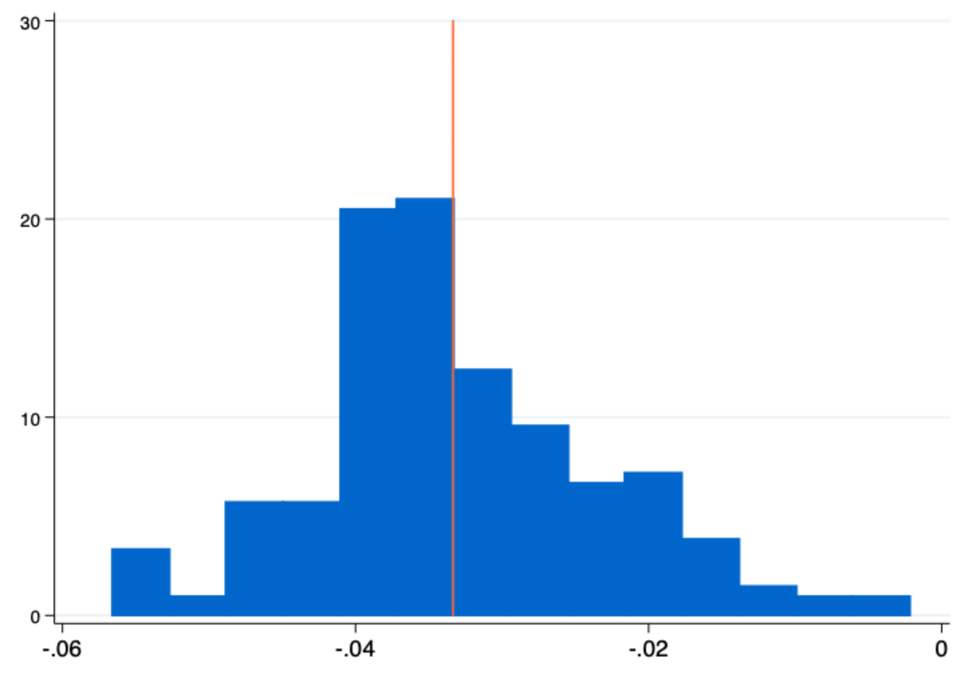
be expected from an operational point of view. The inclusion of that variable in models improved the goodness of fit, as assessed by the normalised RMSE, and did not have significant impact on the sensitivity of modelled costs. On this basis, we have included meter penetration in each of the models within the suite we are putting forward.

Economies of scale

We considered the evidence for including an explanatory variable to capture potential economies of scale. For models where the dependent variable was defined as the logarithm of cost per household basis, we did this by including in the set of explanatory variables a measure of the logarithm of the number of households served. Analogously, in models where the dependent variable was defined on a per service basis, we considered within the set of explanatory variables the logarithm of the number of services provided.

We found that in models with either formulation of the dependent variable, the estimated coefficient on the metric capturing economies of scale were similar: – 0.036 in a model based on a per household specification and – 0.040 in a model based on per service specification. Taken at face value, these estimated coefficients would suggest some economies of scale. The models would predict that a 10% difference between the scale of two companies would be associated with unit costs that are 0.36% / 0.40% lower. However, in the models considered, we found that the estimated coefficient for the variable were not significant at the 10% confidence level. When examining the sensitivity of the estimated coefficient to variations in the dataset, we found that the estimated coefficient on those scale variables covered a relatively wide range though remained almost always negative. Figure 15 illustrates this in the case of a model expressed on a per household basis.

Figure 15 Histogram of estimated coefficient on Ln (Number of households) for model O2



In addition to the statistical evidence, there is an argument that smaller companies have opportunities to mitigate potential cost disadvantages from their smaller size (e.g. through joint ventures or commercial agreements with other companies who bill some of the same customers) and that is it not appropriate from a regulatory perspective to make customers of smaller companies pay extra through allowance for economies of scale in the retail benchmarking models. However, it is also possible that smaller companies may not be able to fully offset scale effects (perhaps due to relative bargaining power).

Overall, in light of the statical evidence and wider economic considerations, we see a mixed picture on whether to include a variable to control for economies of scale explanatory in models comparing costs per household. For this type of models, we followed Ofwat's PR19 approach of including models with and without this explanatory variable in our model suite. In respect of models that benchmark cost per service, we found there too that the variable relating to economies of scale was not statistically significant (at 10%, or even 20%, significance level) and that the inclusion of it in models contributed little or nothing to improving the goodness of fit whilst adding considerably to the sensitivity of modelled costs. For those type of models, i.e. models of cost per service, we opted to not control for economies of scale.

Model dynamics

In developing the models for other retail costs we considered alternative specifications with regard to their dynamics. Specifically, we explored:

- Models that kept the static specification of the PR19 models.
- Models that included a set of year-specific dummy variables, one for all but one of the years spanned by the dataset.
- Models that included a variable capturing a time trend.
- Models that included a time trend as well as three year-specific dummy variables, one for each of the years between 2019/20 and 2021/22. The rationale for including this set of dummy variables is that of controlling for potential Covid-19 effects that lead to costs diverging from the time trend estimated across the whole sample period.

The backdrop to considering alternative dynamic specification of the models is the observed changes over time in companies' other retail costs. This is shown in Figure 16, which charts industry average other retail costs per household over 2013/14 to 2021/22.

Figure 16 Industry-average other retail costs per household, 2013/14 to 2021/22 (£/ household, 2017/18 CPIH adjusted)

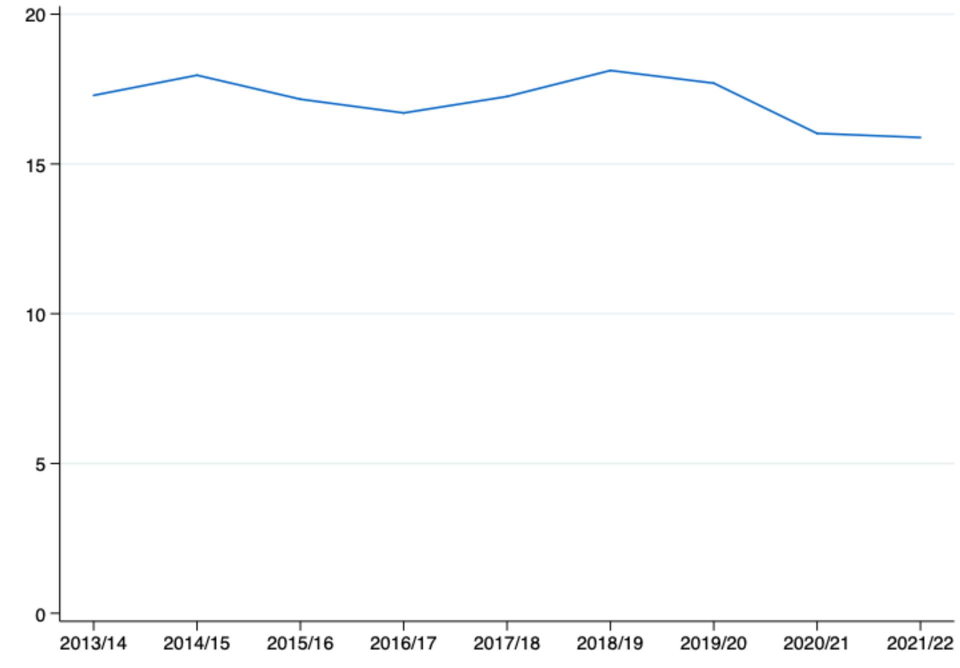


Figure 16 suggests that the industry-average unit cost relating to other retail costs has fallen marginally, in CPIH-adjusted terms, over the period from 2013/14 to 2021/22. One point of interest is that, from the figure alone, it is not evident that there was an impact on industry's other retail costs associated with the Covid-19 pandemic; at least, not one that is as noticeable as is the case for bad debt related costs, which was shown in Figure 8 earlier. We explored this more formally in our analysis by considering, as set out earlier, variants of models that included dummy-variables for each of the last three years in the sample period.

We found that models which allowed for some dynamic effects performed better than ones that, like those at PR19, did not. Such models produced a better goodness of fit. For example, and drawing on model O1 or O3 in Table 10 as examples, the R-squared of a variant of that model which did not allow for dynamics in the specification was 0.12, compared to an R-squared of 0.18 for a variant that included year-specific dummy variables and of 0.15 for one that included a time-trend. In model variants that included some dynamics, we found that the sensitivity of predicted unit costs tended to be greater than that for models with no dynamics, though the difference was relatively small (around 8% in the jackknife metric of sensitivity we drew on).

One aspect of interest is that in variants of the model that did not control for dynamics – like the specifications of the models at PR19 – the variable controlling for variations in meter penetration was not statistically significant and its estimated coefficient was negative, which is counter-intuitive.

As listed earlier, one of the variants we examined was that of models which included both a time trend and a set of year-specific dummy variables for each of the three years from 2019/20 to 2021/22. This was to consider the hypothesis that, over the Covid-19 period, the trend in other retail costs varied from that observed over the earlier years in the sample period. Some of the companies' commentaries in their APR submissions for those years put forward the Covid-19 as a cause of some of the impact on cost lines that are within the 'other retail cost' category, though the direction of such impacts were in both directions (e.g. downward impact on meter reading costs due to reduced frequency in meter reading, upward impact on costs associated with accommodating staff working from home). We had no a priori view on the direction of the net impact of Covid-19 related factors on other retail costs or indeed whether the net impact was material across the industry. Considering a variant that included year-specific dummy variables for those Covid-19 years, in addition to a time-trend, allowed us to explore that question.

Table 11 shows the regression results for models O3 and O4 when these are estimated with (i) a time-trend, and with (ii) a time trend plus dummy-variables for each of the last three years of the sample period, i.e. for 2019/20, 2020/21 and 2021/22. The table reports the estimated coefficients and below each of these, enclosed within brackets, their t-statistic.

Table 11 Regression results for variants of selected models of other retail costs

	Model O3		Model O4	
Variant on dynamics	Time trend	Time trend plus Covid-19 dummies	Time trend	Time trend plus Covid-19 dummies
Explanatory variables				
Proportion of dual service households	0.210 (2.395)	0.210 (2.421)	0.292 (3.562)	0.291 (3.713)
Proportion of metered households	0.555 (2.045)	0.490 (1.729)	0.577 (2.163)	0.505 (1.822)
Logarithm of number of households			-0.035 (-1.095)	-0.035 (-1.118)
Time-trend	-0.020 (-1.945)	-0.006 (-0.478)	-0.020 (-1.951)	-0.006 (-0.483)
Dummy variable for 2019/20		0.005 (0.147)		0.005 (0.142)
Dummy variable for 2020/21		-0.090 (-2.045)		-0.089 (-2.037)
Dummy variable for 2021/22		-0.114 (-1.715)		-0.114 (-1.709)

R-squared	0.15	0.17	0.16	0.19
Observations	153	153	153	153

An observation from the regression results presented in Table 11 is the drop in statistical of some the independent variables when moving from models with a time trend only to ones with a time trend plus year specific dummy-variables for the Covid-19 period. This is particularly the case for the time-trend variable itself. We also note that the t-statistics on the year-specific dummy variables for 2019/20 is very low and that for 2021/22 it is also below 2, in absolute terms, and that that variable is not statistically significant at the 10% level. Finally, whilst the R-squared of the model variant that included the set of three dummy variables is, necessarily greater, the difference is somewhat marginal and on our goodness of fit metric, the normalised RMSE, the two variants perform very similarly.

On the back of the analysis carried out within this project, we considered that, unlike the models of bad debt related costs, there was not strong evidence to prefer a more complex structure which controlled for Covid-19 effects, through the inclusion of year-specific time dummies in each of the last three years of the sample period. Once further data are available, for 2022/23 and subsequent years, it may be useful to revisit this question.

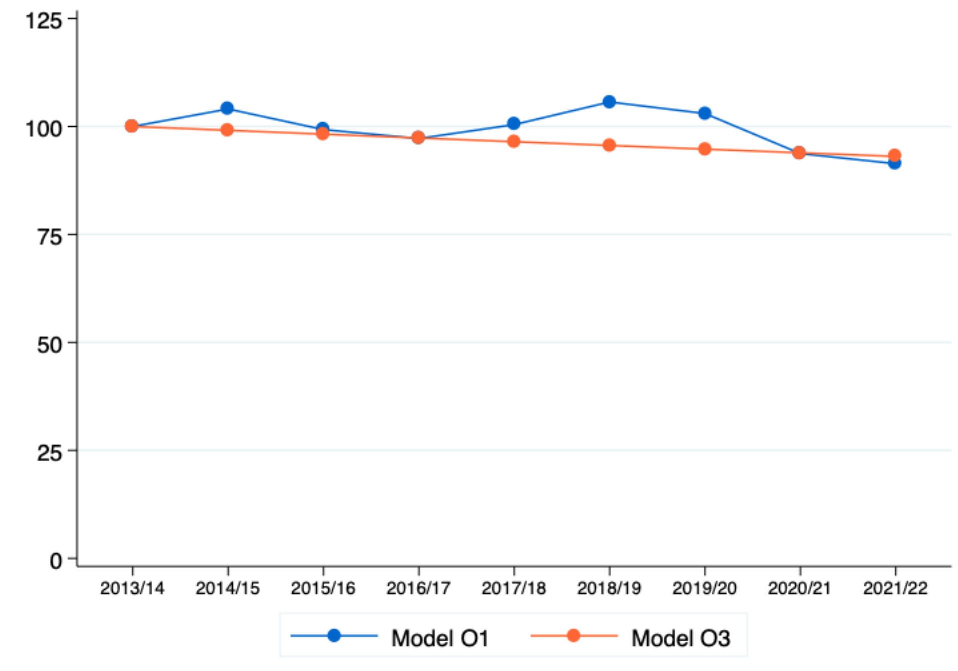
In that light, we opted to put forward models of other retail costs per household and per service containing either (i) year specific dummies for all years in the dataset or (ii) a time trend.

For those variants of the model that include a time-trend – models O3, O4 and O6 in Table 10 – we can read the estimated time trend from the coefficient of the time-trend variable.[12] As such, all those three models predict a downward time-trend over the sample period in other retail costs per household, or per service, of around –2%. This can be interpreted as indicating a year-on-year decrease in modelled other retail costs per household, or per service, of around 2%, in real terms (having controlled for changes to cost driver explanatory variables).

It is less immediate to read the trend over time that is implied by the set of models that include year-specific dummies, models O1, O2 and O5 in Table 10. Those models do, however, predict a similar time profile of unit costs as the models that include a time trend. This is shown in Figure 17 where we contrast the time-trend of unit costs predicted by model O1 and the time-profile of unit costs predicted by model O3; the two models are identical except for the choice of how the dynamics are specified. To draw the figure, we controlled for the effects of the other relevant variables, with a view to isolating the effect associated with time. For ease of interpreting the figure, we have rebased, for each model, the predicted unit costs for 2013/14 at 100.

[12] Given the functional form of the models, the predicted time-trend is estimated by taking the exponential of the estimated coefficient on the time-trend variable. Across the relevant set of models, that estimated coefficient is around -0.02. The exponential of such a small value can be approximated by (1-0.02) = 0.98, which points to a 2% decreasing trend.

Figure 17 Trend in real (CPIH-adjusted) other retail costs per household implied by models O1 and O3 (2013/14 set to 100)



As shown in the figure, the time-profile predicted by the two types of models follow each other relatively closely, although with some divergence from 2017/18 to 2019/20. We have included both variants within the set of models we are putting forward in this study, as set out in Table 10.

6: Disaggregated versus aggregated models

At PR19, for its residential retail cost assessment, Ofwat drew on a suite of models that included a set of disaggregate cost models – models that focused separately on bad debt related costs and on other retail costs – and on a set of aggregate cost models relating to total retail costs.

In its triangulation of model results, for the purpose of determining companies' modelled costs, Ofwat gave 75% of weight to the modelled costs derived from the aggregate cost models and 25% to the modelled costs calculated from summing the results derived from the models of bad debt costs and of the models of other retail costs.

In the context of residential retail cost assessment, Ofwat had signalled in its PR24 Draft Methodology its intention to focus on models of total retail cost and to not develop bottom-up estimates. Subsequently, in its PR24 Final Methodology, Ofwat appears to have back-tracked from that position and set out a less definitive position with regard to how it will draw on disaggregate and aggregate cost models. We reviewed Ofwat's reasoning on this earlier in section 3.

In this section we consider the case for models of total retail costs. We take the following points in turn:

- We set out conceptual motivations and risks in drawing on models of total costs in the context of household retail.
- We set out candidate models of total retail costs.
- We set out analysis on the performance of models of total costs relative to models of disaggregated costs.

Conceptual motivations and risks in models of total cost

Especially where the sample size is relatively small, there are some potentially important benefits from using models of more disaggregated cost categories rather than models of aggregated costs. In particular:

- Where a given cost driver affects (or primarily affects) only a subcategory of costs, then using a benchmarking model that is targeted on that subcategory of costs provides a way to focus the model estimation on the relationship between that cost driver and those costs. In contrast, if modelling is carried out for a broader category of costs, the estimation of the relationship between the cost driver and costs may be polluted and by the model trying to associate that cost driver with other categories of costs it has little or no impact on.
- Given the small size of the dataset available for benchmarking water companies' retail costs, it may not be desirable or feasible to include within the set of explanatory variables of the aggregate cost models all those drivers which are found to be important in each of the disaggregated models. We saw at PR19 that Ofwat included a narrower set of cost drivers in its aggregated models of residential retail costs compared to its disaggregated models. Even where cost drivers from disaggregated model are included in aggregated models, the larger set of explanatory variables, and less direct relationships with costs, may mean that estimated coefficients are subject to greater estimation uncertainty (e.g. as indicted by lower t-statistics for

some coefficients or variables that were deemed statistically significant in disaggregated models not being deemed statistically significant in aggregated models) – again, we saw this in Ofwat's PR19 model suite.

These considerations are particularly relevant where there is little, or no, overlap in the cost drivers relevant to each of the disaggregated categories of cost. Were it the case that the same set of cost drivers were relevant to bad debt related costs and to other retail costs, then a model of aggregate retail cost would potentially not need to be expanded to cover a greater number of variables than those already captured in each of the disaggregated models. In the event, however, that is not the case: the cost drivers in each of the two disaggregated cost models do not overlap.

The above considerations point to there being potential downsides of aggregated models compared to disaggregated cost models. But there are also potential factors that may mean that an aggregate cost model can make a useful contribution to the cost assessment, whether alongside or in place of disaggregated models. In particular, the following two factors could in principle have the effect of worsening the accuracy and reliability of disaggregated models, without causing corresponding problems in aggregated models:

- Differences between companies in their approach to allocating their costs between the disaggregated cost categories.
- Where companies' choices of how to run and organise their business have a significant impact on how costs are distributed across the disaggregated cost categories.

Ofwat referred to these points in its consultation on econometric cost modelling for PR19 where it commented that the disaggregated models it had selected – models of bad debt costs and models of other retail costs – were not immune to cost allocation issues and to potential trade-offs between them.[13]

At the conceptual level, there are reasons in favour of using disaggregated models and reasons for using aggregated models. Nonetheless, in any given case, we see a role for carrying our empirical analysis to look for evidence that may inform on the relative merits of disaggregated versus disaggregated modelling results.

For instance, if concerns about the consistency of cost allocation and internal organisation across companies had an important influence on the cost benchmarking results for disaggregated models, without offsetting benefits from disaggregated modelling, we would expect to see the effects of this come across in the performance of disaggregated models compared to the performance of aggregated cost models.

The analysis set out in the remainder of this section is informative of how, in the specific context of models to benchmark residential retail costs, the points raised above on the relative merits of disaggregated and aggregated models, play out in practice.

[13] Ofwat (2018) *Cost assessment for PR19: a consultation on econometric cost modelling*, page 23.

Candidate models of total retail costs

We explored a series of alternative models to benchmark companies' total (residential) retail costs. We developed alternative candidate models of total retail costs by drawing on the same set of cost drivers that we had considered when developing models of bad debt related costs and models of other retail costs. Out of the range of total cost models we explored, we report in Table 12 those which performed better in terms of goodness of fit, variance of predicted unit costs and in terms of the estimated coefficients seeming to be reasonable from an operational / economic point of view.

Table 12 Econometric models for total retail costs

Model ref	T1	T2	T3	T4
Dependent variable	Ln (Total retail costs per household)			
Explanatory variables				
Ln (revenue per household	0.422*** (3.835)	0.402*** (3.805)	0.441*** (4.441)	0.426*** (5.029)
Credit risk score (eq_rgc102)		–0.018 (–1.534)		–0.014** (–2.102)
Income deprivation score (unadjusted)	0.014 (0.875)		0.012 (0.787)	
Proportion of metered customers	0.004 (0.923)	0.005 (1.100)	0.003 (0.907)	0.004 (1.118)
Time trend			–0.019* (–1.828)	–0.022** (–2.005)
Dynamics	Dummy variables	Dummy variables	Trend plus 2020–2022 dummies	Trend plus 2020–2022 dummies
Overall R-squared	0.61	0.61	0.61	0.61
Observations	153	153	153	153

We highlight the following points on the models and results presented in the table:

- Cost drivers related to proportion of dual service customers, to scale of company or to mix of services – which we have included in some of the disaggregated models of other retail costs set out in the previous section – did not perform well in models of total retail costs. The variable related to revenue per customer, which we do include in the total cost models, is strongly correlated to the proportion of dual service customers and so is likely to be explaining some of the variation in costs which, in the disaggregated model for other retail costs, was being controlled for by the dual service variable.
- Measures of deprivation or arrears risk and the driver capturing the average revenue per household explain some of the variation in total retail costs per household. This is in line with our finding about their role in explaining variations in companies' bad debt costs. However, the relative confidence in the estimated coefficients on the measures of deprivation or arears risk is reduced from that which we found for models of bad debt costs, i.e. the t-statistic on those

variables are lower than is the case for the models of bad debt costs. In terms of the impact on predicted costs, this lower significance will tend to increase the variance of the unit costs predicted by the total retail cost models.

- The t- statistic for meter penetration is considerable lower for the models of total retail costs than for the models of other retail costs.
- There are no explanatory variables for which the t- statistic is significantly higher in the total cost models than in the relevant disaggregated model. This casts doubt on the idea that the total cost model brings benefits to the accuracy of the modelling.

Comparing aggregate and disaggregate cost models

We have carried out further analysis to compare the relative performance of the models of total cost against that of the set of disaggregated models for costs of bad debt and for other retail costs. In particular, we have sought to assess how the predicted values obtained from combining the results of two disaggregated models – one for bad debt related costs and one for other retail costs – compare with the predicted values of an aggregate cost model in respect of goodness of fit and in terms of the variance of the predicted costs.

For this analysis, we contrasted the four combinations of disaggregated and aggregated cost models set out in Table 13.

Table 13 Grouping of models considered in analysis of disaggregated vs aggregate cost models

Contrast ref.	Bad debt related cost model	Other retail cost model	Total retail cost model
C1	D2	O1	T2
C2	D1	O1	T1
C3	D4	O3	T3
C4	D5	O3	T4

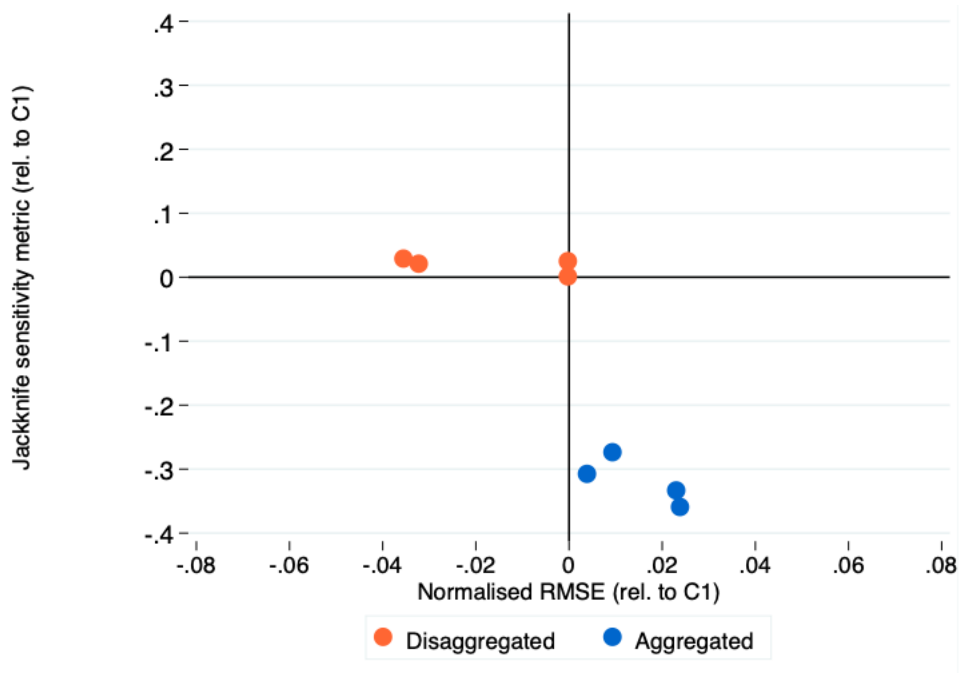
The analysis related to contrast C1 for example, involved comparing the performance of the results obtained from bringing together the results of disaggregated models D2 and O1, with the results of the total retail cost model T2. We defined the set of contrasts shown in the table such that, in each of them, the total cost model and the disaggregated cost models shared similar features with regard to the specification of the dynamics and the specification of the dependent variable (we did not specify contrasts that drew on models of bad debt related costs where the dependent variable was expressed as the ratio of cost to revenue) and to the measure of deprivation or arrears risk included. This was aimed at obtaining a clearer comparison between each of the pairs of disaggregated models and the total cost model.

For our analysis:

- For each of the contrasts specified in Table 13 we calculated (i) the predicted total unit cost derived from the two disaggregated models, and the (ii) the predicted total unit cost derived from the total cost model.
- Drawing on those values, we calculated the normalised RMSE in respect of the "disaggregated predicted unit total cost" and in respect of the "aggregated predicted unit total costs".
- For each of "disaggregated predicted unit total cost" and "aggregated predicted total unit cost" we calculated a measure of the predicted unit cost variance, based on the jack-knife-like approach we presented in section 2 and describe in more detail in Appendix 1. Those metrics captures the variance around the two different sets of predicted unit total costs.

Figure 18 presents the outcome of the analysis. The figure contrasts the relative goodness of fit and the jackknife sensitivity metric of unit total retail cost for disaggregated cost model and aggregate cost models for the grouping of models presented in Table 13. For the purpose of drawing the figure, we have normalised figures using the results for the disaggregated models under contrast C1 as a reference point, so that the metric on jackknife sensitivity and on normalised RMSE of models have been expressed relative to the value of those metrics for contrast C1.

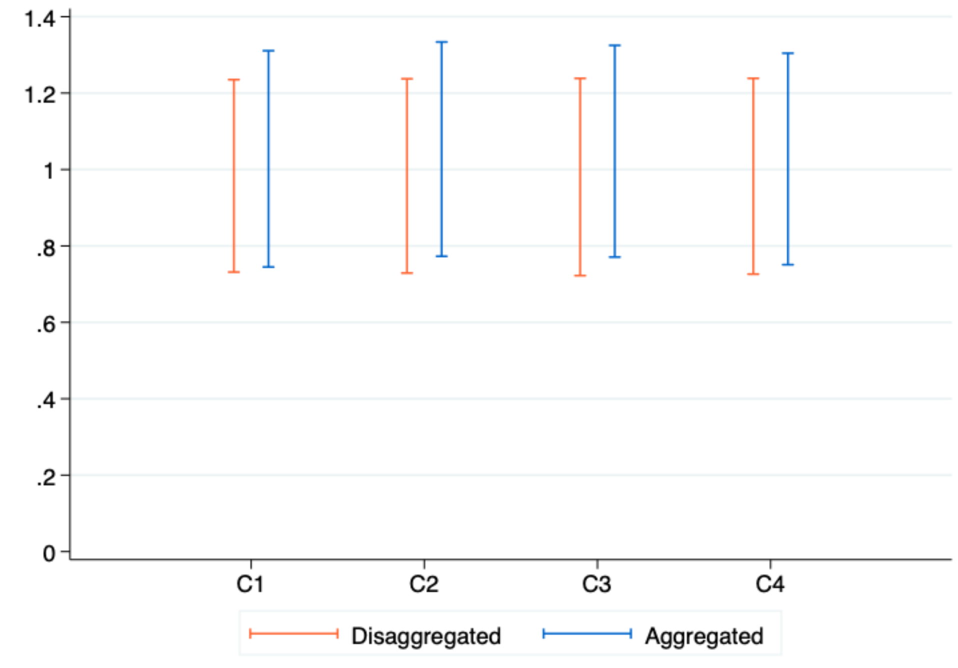
Figure 18 Comparison of performance of disaggregate and aggregate cost models



As shown in the figure, the combination of pairs of disaggregated models perform slightly less well with respect to goodness of fit and considerably better with respect to the sensitivity of modelled unit (total) cost. The results of this analysis echo the points raised earlier, in relation to the loss of statistical significance of some of the variables in the total cost models compared to their role in disaggregated models.

Figure 19 shows the spread of the efficiency ratios – the range from the minimum to the maximum value of the efficiency ratios across companies – for the aggregate and (the pairs) of disaggregate cost models that we set out earlier. The ratios shown in the figure are based on companies' average efficiency ratios over the period 2017/18 to 2021/22. The spread of these efficiency ratios is one other indicator associated with the goodness of fit of models and, as suggested in the figure, the range tends to be narrower when disaggregated cost models are considered, though the difference is not great.

Figure 19 Spread of efficiency ratios for disaggregate and aggregate cost models



Our analysis of the performance of models of total cost suggests that there is no evidence-base to favour these over disaggregated cost models; rather, the evidence points in the opposite direction. As set out here, the analysis required going a bit beyond the more conventional metrics drawn on to assess and compare model performance, though the finding is also visible from the more conventional metrics drawn on to review models, namely the statistical significance or t statistics of explanatory drivers.

For the purposes of this report, we focused our suite of models on disaggregated models only. While we would not completely rule out the use of aggregated models for residential retail costs, there seemed a weak case overall in their favour. Furthermore, there are benefits from omitting aggregated models as a means to keep the suite of models manageable while also allowing for variation to be retained where it seems to be more important (e.g. in terms of the range of cost drivers considered and alternative approaches for model dynamics).

Appendix 1: Enhanced metrics and techniques to support model assessment

In this appendix we provide more detailed information on some enhanced metrics and analytical techniques that we used as part of our model development and review process.

We take the following in turn:

- Analysis of the sensitivity of estimated coefficients to variations in dataset.
- Tailored metric of goodness of fit.
- Tailored metric of the statistical sensitivity of estimated cost benchmarks.

In the final subsection of this appendix, we discuss how we bring together the second and third elements above into a chart which we consider particularly useful for comparing the performance of alternative model specifications.

Analysis of the sensitivity of estimated coefficients to variations in dataset

One tool we draw on as part of our assessment of models is to examine the sensitivity of estimated coefficients to variations in the dataset that is used for the estimation. This is particularly useful when looking to assess whether the inclusion of specific explanatory variables in models are producing results that might be spurious or misleading, especially in the context of econometric models estimated on a relatively small data sample.

Such analysis can provide insight into the relative robustness of modelling results and be of additional value compared to an examination of t-statistics alone. This is especially so in a context where the number of observations on which regressions are run is limited and where there are likely to be correlations between residuals (e.g. over time for the same company).

The mechanics of the approach is somewhat brute force in nature: we consider variations in the dataset and, for each variation of the dataset, we estimate the model being considered and extract the set of estimated coefficients for each of the explanatory variables.

We take a systematic approach to varying the dataset. Specifically, we produced variations of the dataset by:

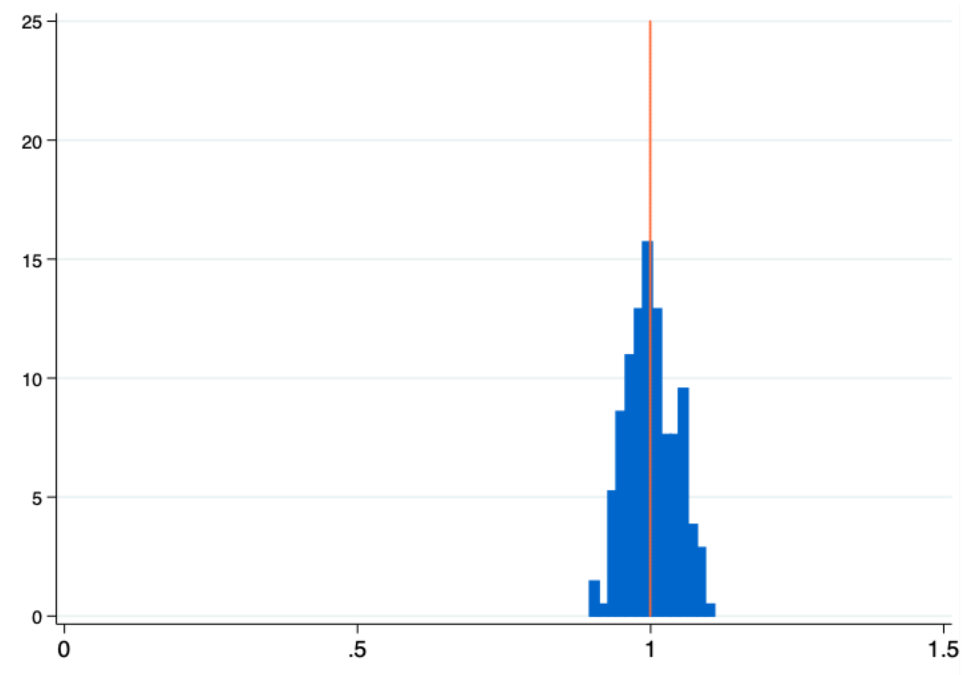
- For each year at a time, dropping the observations for that year.
- For each company at a time, dropping observations for that company.
- For every combination of a single year and a single company, dropping observations that refer to either that year or that company, or both.

In the context of the dataset relevant to retail cost assessment covering the period 2013/14 to 2021/22 this produced 210 variations of the dataset (including the variation with all observations).

Beyond the exact algorithm we used to produce the variations to the dataset, the value in the approach that we would emphasise is the overall method of creating tweaked versions of dataset and analysis the sensitivity of estimation results to this. There are choices around how exactly to create the set of tweaked datasets (e.g. how many companies to drop at a time, whether to drop companies and years simultaneously) and, in turn, how many datasets get created.

A useful means to visualise the results from such sensitivity analysis is to draw a histogram of the estimated coefficient for a given variable. Figure 20 shows an example of this. The chart relates to the estimated coefficient for the logarithm of average bill in one of the bad debt related cost models that we explored, namely model D2 which we presented earlier in section 4. The vertical orange line marks the value of the estimated coefficient when the full dataset is used.

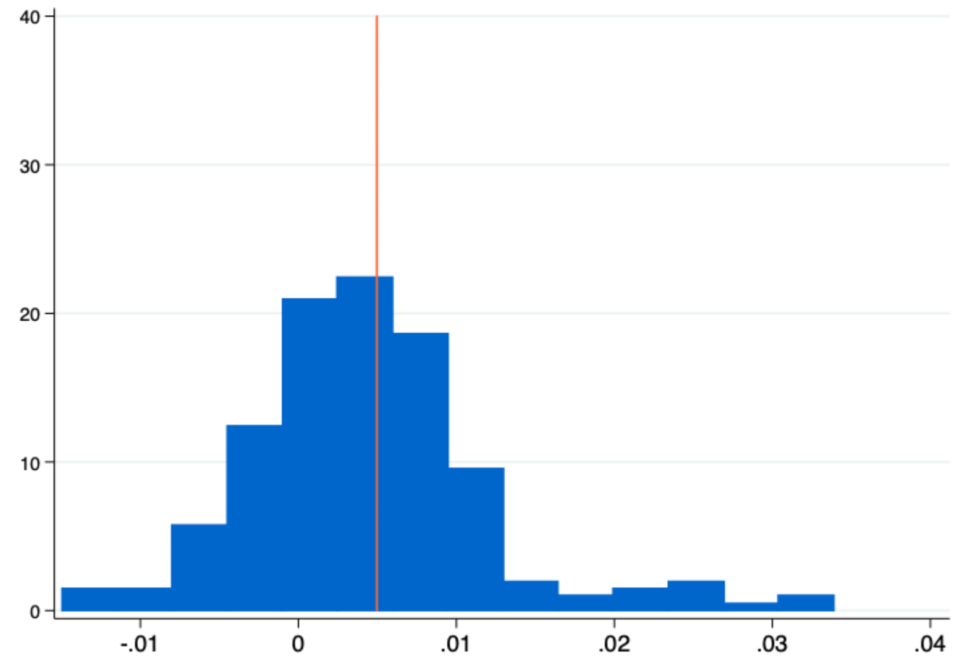
Figure 20 Histogram of estimated coefficient on logarithm of average bill in model D2



The histogram shows that, across the variations in the dataset, the estimated coefficient for the variable is within a relatively narrow band, and that this is far removed from the value of 0. On that basis, the histogram suggests that the estimated coefficient on the explanatory variable related to the size of the average bill is robust to small variations in the dataset underlying the estimation.

In contrast to the above, Figure 21 shows the histogram for the estimated coefficient for the variable related to the customer transiency in one of the models related to bad debt costs that we explored, and which we discussed in section 4.

Figure 21 Histogram of estimated coefficient on metric on household transiency



As shown in the figure, the histogram for the estimated coefficient for the variable proxying household transiency is relatively wide and straddles the value 0. Whilst we would not take this histogram in isolation to form a view on the appropriateness or not of including such a variable in the modelling, the figure is informative, and this type of analysis has formed part of our assessment of models and modelling approaches.

Tailored metric of goodness of fit

Goodness of fit is an important aspect of model performance. At PR19, the main metric used for this by Ofwat is R-squared. There are two key limitations or concerns in the context of the retail cost benchmarking:

- It is misleading to compare R-squared between models with different dependent variables (e.g. between models in which the dependent variable is an aggregate cost measure and models that use a unit cost measure for the dependent variable, or between unit cost models which involve different measures of units). It is quite possible that an aggregate cost model that has a much higher R-squared than a unit cost model is nonetheless a worse model in terms of its ability to explain the variation in underlying costs between companies.

- Adding more explanatory variables to a model will inevitably increase the R-squared, even if the relationships between those additional explanatory variables and the dependent variable are coincidental or spurious. So, it can be difficult to use R-squared to compare the performance of models which have different sets of explanatory variables. Some sources suggest the use of an "adjusted" R-squared metric, but this does not seem an altogether reliable solution: the derivation of the adjusted R-squared metric involves, in effect, applying a penalty to the R-squared measure which varies according to the number of explanatory variables in the model, without a sound basis for the scale of the penalty.

Ofwat is alive to the first of these concerns. Its recent guidance on the template to use for the submission of cost models emphasis that the R-squared measure, adjusted or not, should not be used to compare models where the dependent variable is defined differently.[14] At the same time, Ofwat seems to conflate the issue of the magnitude of the R-squared measure with the question of how much of the observed variation in companies' costs a model can explain: the matters are related, but distinct.[15] The R-squared metrics are concerned with how well the model can explain the observed variation in the dependent variable that is being modelled, e.g. the variation in the logarithm of cost per household, rather than the variation in companies' costs in £ million.

We have sought to deal with these issues.

The first point highlighted above is a symptom of how the R-squared metric itself is constructed. Our approach has been to develop an alternative metric which can be interpreted as a normalised goodness of fit measure which is comparable across models where the specifications of the dependent variable differ though they relate to the same element of cost that is being benchmarked. For example, it allows for a comparison of a model where the dependent variable is expressed as the logarithm of bad debt related costs per household with a model where the dependent variable is expressed as ratio of bad debt related costs to average bill. For a given model, this metric is calculated as follows:

1. Use the predicted values from the econometric model estimation results to calculate, for each company and for each year of the data period used in the modelling, the implied modelled costs per household per year (in £/household).
2. Calculate the ratio of the actual costs per household, for each company and year, and the modelled unit values from step (1) above. This provides a set of 'efficiency ratios', derived from the relevant model for that company, for that year.
3. Subtract one from the efficiency ratio and take the square of the result.
4. Take the average, across all companies and years, of the squared normalised ratios from step (3).
5. Take the square root of the figure from step (4).

This is, in effect, a form of **normalised root mean square error (RMSE) metric**, tailored to the retail cost benchmarking exercise. It captures the scale of residuals – and range of implied "efficiency scores" across companies and over time – and condenses this into a single metric that can be compared between models. In this report, for convenience, we refer to this metric as the normalised RMSE.

---

[14] Ofwat (2022) "Template and guidance for the submission of base econometric cost models ahead of the spring 2023 consultation", page 16.

[15] Ofwat (2022) "Template and guidance for the submission of base econometric cost models ahead of the spring 2023 consultation". On page 16, in the context of a discussion on R-squared measures, Ofwat comments that "If a model failed to explain a significant share of the costs of the industry, it would be inappropriate to use it for the estimation of costs. But equally, a strategy for searching for a model with a high R-squared has the risk of finding a model that fits the data well but is in fact incorrect."

Tailored metric of the statistical sensitivity of estimated cost benchmarks

The ultimate role of the econometric benchmarking models is to produce cost benchmarks for each company to be used in setting price controls, based on the predicted values from the estimation of the econometric models. In this context, we consider that it is important – and illuminating – to consider the "precision" with which the modelled costs calculated from the predicted values from the models are estimated. This is something that seems to have been given relatively limited attention to date in Ofwat's econometric previous modelling of water company costs, and in its recent guidance on model submissions for PR24.

There are different ways in which such sensitivity might be gauged. For instance, as a starting point:

- Software packages such as Stata can generate, as part of model post-estimation results, statistical estimates of the variance or standard error of the predicted values from the regression. The estimated variance of each predicted value will be a function of the variance estimated for each of the explanatory variables in the model, and of the value of each explanatory variable *for that observation*. This means that the estimated variance associated with the predicted value for Anglian Water in 2019/20, say, will be different to the estimated variance for Severn Trent in 2019/20 or the estimated variance for Anglian Water in 2018/19
- We can assess how sensitive the modelled costs for each company are to minor variations in the dataset used (e.g. re-running the regression, dropping one company or year from the sample).

We make most use of the second type of approach. This represents an alternative way to gauge the sensitivity of modelled, which has some advantages in a small sample (where the assumptions underpinning the measures of precision that are reported by default within the output of regression results produced by software packages such as Stata might not hold).

Our approach involved testing the sensitivity of the predicted values derived from a given model to minor variations in the dataset. We constructed this sensitivity analysis by running the models on different variations of the dataset and, on each run, calculate the modelled (unit) cost of each company in each of the years covered by the dataset.

The approach we took to produce the variations in the dataset is the same as that we outlined earlier, when we set out our approach to deriving the histograms for the estimated coefficients of explanatory variables. We estimated the model being considered using each of those variations in the dataset, and used the results of these to calculate a measure of the sensitivity of estimated unit modelled costs as follows:

1. For each of the variations in the dataset, we drew on the estimated model results to calculate the predicted (the fitted) values of the model. What these predicted values relate to will differ across models, according to how the dependent variable is specified. For example, the predicted values that are derived directly from the model results may relate to the logarithm of unit costs, to unit costs, to logarithm of aggregate costs or some other cost metric.

2. We used the predicted value in (1) to calculate modelled costs on a cost per household basis, for each company in each year, and across each of the dataset variations. For a given model, these measures of cost per customer differ across companies, years and dataset variations.
3. We normalised the modelled unit costs from step (2), by (a) dividing the modelled cost per household for each specific company in each year for a given dataset variation by the modelled unit cost per household for that company in that year when the full dataset was used. This step puts the figures from step (2) on a common basis.
4. We calculated the statistical standard deviation across the normalised modelled unit costs from step (3).
5. We multiplied the standard deviation from step (4) by the average of the unit modelled cost across companies and years when the full dataset is drawn on, so that the metric is something that is expressed in £ terms and so is easier to interpret.

This method involves something which might be described as a specific form of jackknife resampling in the statistic and econometric literature.

This process produced a single metric, which is expressed in the units of costs per household. It is a measure of the statistical variance of measures of the modelled unit costs from the model estimation process, which have been normalised across years and across companies.

We refer to this metric as the **jackknife sensitivity metric**.

This metric enables direct comparisons between models in the sensitivity of the cost benchmarks predicted by each model. The higher the metric, the greater the variance – or sensitivity – in the predicted costs from the model estimation process to minor changes in the dataset used. All else equal, a model with a lower sensitivity metric performs better and would be preferred.

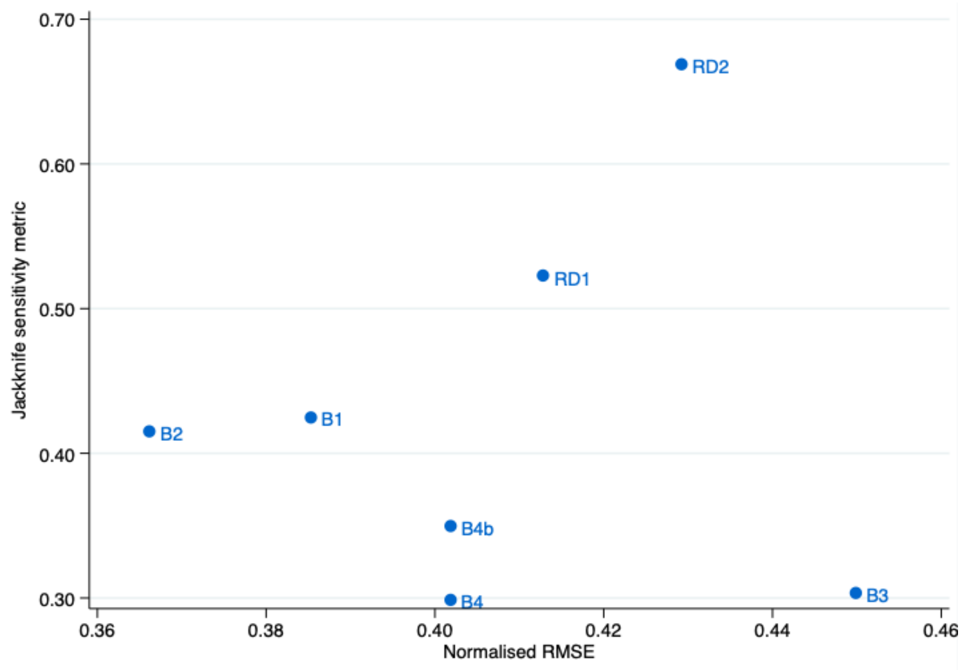
Bringing goodness of fit and sensitivity together

Our jackknife sensitivity metric captures the performance of a model in terms of the relative sensitivity of the modelled unit costs. It is not a measure of goodness of fit. Indeed, in developing and choosing between alternative econometric model specifications, there may be a trade-off between goodness of fit and sensitivity of modelled costs.

- Adding additional explanatory variables to a model may increase goodness of fit, but in some cases, this may lead to substantial increase in the sensitivity of modelled costs, which might be seen to lead to a worse model overall.
- A model with only a single explanatory variable might appear to perform well in terms of sensitivity compared to a model with several variables because it does not allow as much scope for estimation error to creep in (few parameters are estimated via the regression). But that model might have limited explanatory power and relatively poor goodness of fit.

In this context, we find it helpful to produce scatter charts which show how different models perform in the two important dimensions of goodness of fit (measured by our normalised RMSE metric) and statistical sensitivity of the estimates of modelled costs (measured by our jackknife sensitivity metric). Figure 22 illustrates this for a set of models of bad debt costs.

Figure 22 Jackknife sensitivity metric vs normalised RMSE for selected models of bad debt costs



Models RD1 and RD2 refer to the two bad debt related cost models that Ofwat drew on at PR19 (estimated, however, on a dataset spanning 2013/14 to 2021/22). The other models are ones we have explored in the context of this work, and which, with the exception of the model labelled B4b, are presented in more detail in the section below. The figure shows, for example, that Ofwat's PR19 model RD2 performs distinctly less well on both the jaccknife sensitivity metric of modelled costs and on the goodness of fit measure than the remaining models pictured. We make two other comments on the figure. First, as shown in the chart, there will be comparisons between models where there is a trade-off between their goodness of fit and the sensitivity of the predicted modelled costs that are derived from them. That is the case, for example, where one to be comparing between models B2, B3 and B4. Second, we suggest that the type of chart set out above can be useful when assessing and comparing variants of particular models. Take the models labelled B4 and B4b; the latter is a variant of the former in that its set of explanatory variables includes, in addition to those variables included in model B4, a measure capturing customer transiency. Observing the relative performance of these two models in the figure, it can be seen that including the additional explanatory variable has very limited impact in improving the goodness of fit but adds significantly to the sensitivity of modelled unit costs.

Appendix 2: Further information on data used

This appendix sets out further details on the data used in the analysis.

Data sources

Table 14 sets out the sources of data we drew on.

Table 14 Data sources

Source	Comment
Ofwat Oct 2022	We extracted data from worksheet labelled "nom statafile" of the Excel file "PR24_Cost_Assessment_Master_Dataset_Residential_Retail_v3.0.xlsx", which Ofwat published in November 2022.
ONS IMD2019	We derived measures of the square of income deprivation score related to population served by water companies drawing on the data on England and Wales comparable measures of income deprivation score, published by the ONS in context of the IMD 2019.

Information on derivation of variables

For much of the data manipulation and derivation of relevant variables, we have drawn on the Stata code set out in the file "Residential retail do file v2.0.do", published by Ofwat in November 2022. In broad terms, we have followed the set of assumptions regarding how the different variables relating to costs and to candidate cost drivers are defined and constructed, and we have followed the approach coded in that script, and spelt out in the accompanying guidance document, with regards to the treatment of companies that merged over the period covered by the data.[16] We do not repeat those common aspects here. In Table 4 we highlight instead where we have departed from Ofwat's approach or where the code in that Ofwat Stata file constructed alternative measures and did not obviously settle on a particular one. We highlight too information on variables which we drew on in our analysis and which are not derived in Ofwat's Stata script.

Table 15 Aspects relating to derivation of data used in analysis

Aspect/variable	Comment
Depreciation	We used outturn, rather than smoothed, depreciation in constructing relevant measures of costs used in analysis.
Doubtful debts	We used the measure of "smoothed" doubtful debt to derive the variables on bad debt related costs.

[16] Ofwat (2022) "Template and guidance for the submission of base econometric cost models ahead of the spring 2023 consultation".

Aspect/variable	Comment
Revenue	We adjusted the revenue figures reported in the dataset to deal with the £50/customer contribution from Government that South West Water customers benefit from. Specifically, we made the following adjustments to the revenue figures for SWT and for SWB: • For SWT we deducted £50 per customer. • For SWB, (i) for each of the years from 2013/14 to 2015/16, we deducted (£50 * Number of SWT customers); (ii) for each of the years from 2016/17, we deducted (£50* Number of SWB customers* *ratio*), where *ratio* is the average of the ratio of SWT customers to the sum of SWT and BWH customers in the years from 2013/14 to 2015/16.
CPIH-adjusted values	We used CPIH-adjusted measures of cost and revenue in the definition of relevant dependent and explanatory variables.
Square of ONS income deprivation score	We derived a metric capturing the company-level aggregate of the square of LSOA-level income deprivation scores published by ONS in the context of the IMD2019. We extracted the underlying LSOA-level data used from the dataset published by ONS which puts English and Welsh LSOAs on a comparable basis. Drawing on that data, rather than the data on the income deprivation scores published separately for England and for Welsh areas, addresses the concern around the differences in the set of criteria used by the ONS and by Statistics for Wales when deriving those statistics for their nations alone. A discussion on the comparability issue is set out here: https://www.gov.uk/government/statistics/indices-of-deprivation-2019-income-and-employment-domains-combined-for-england-and-wales/indices-of-deprivation-2019-income-and-employment-domains-combined-for-england-and-wales-guidance-note The LSOA-data we drew on is available from that link too. Such comparable data are available only for the income scores published in the context of the 2019 IMD. We have applied data from that for all years spanned by our sample period. To aggregate across LSOAs within company service areas, we used a mapping of LSOAs to waster and to sewerage service areas which we produced by drawing on geographical data on the boundaries of LSOAs and of companies' service areas.

Appendix 3: Results from re-running PR19 models

Introduction

This appendix presents the results for the set of econometric models Ofwat drew on at PR19, updated to reflect the additional that have become available since.

We take Ofwat's PR19 models of bad debt related costs, models of other retail costs and models of total retail costs and for each present:

- the model results as put forward by Ofwat at PR19, to be seen as a reference point;
- the model results based on the dataset that Ofwat published recently, in Autumn 2022 but using the same time period as in the original PR19 models, i.e. based on data for 2013/14 to 2018/19.
- as the previous case but using the full time period covered by the recent dataset, i.e. using data for 2013/14 to 2021/22.

Differences between the first and second set of results will reflect revisions to the data that Ofwat has done since PR19. We expect these to be small, but there are some. For example, since publishing the PR19 models Ofwat obtained updated data from water companies on net recharges, which are used in the new dataset. In addition to this, in the recent dataset, smoothed depreciation, which Ofwat calculates, and which is one of the components of other retail costs, and of total retail costs, is calculated by smoothing over a longer time period (2013/14 to 2021/22) than was the case in the dataset used at PR19.

We have followed Ofwat's naming approach for models in all tables as provided at PR19. For example, RDC1 refers to Ofwat's first bad debt plus debt management cost model at PR19. Under the heading of each of the models, we report on the three variants, outlined above and which we label with: "As at PR19"; "Data to 2019"; and "Data to 2022".

Bad debt related cost models

Table 16 sets out the model results for bad debt related cost models, drawing on the structure outlined above.

Table 16 Variants of Ofwat's PR19 bad debt cost models (RDC)

Model Ref	RDC1			RDC2		
Model variant	As at PR19	Data to 2019	Data to 2022	As at PR19	Data to 2019	Data to 2022
Dependent variable	Ln (Bad debt costs per household)					

Explanatory variables						
Total household revenue per customer	1.190***	1.129*** (11.434)	1.188*** (9.312)	1.158***	1.118*** (12.211)	1.164*** (9.536)
Percentage of households in default	0.067***	0.066** (2.348)	0.024 (1.256)			
Income score (unadjusted)				0.076***	0.054* (1.878)	0.021 (0.856)
Customer transience/migration				0.035**	-0.001 (-0.035	-0.015 (-0.651)
Overall R-squared	0.77	0.75	0.62	0.78	0.74	0.61
Observations	105	102	153	105	102	153

*Note: For each explanatory variable included in a model, the table reports: value of estimated coefficient; in brackets, the z-statistics; and whether the coefficient is significant at the 1, 5 or 10 per cent significance level, indicated by ***, ** and * respectively.*

Some comments on the table:

- For the model RDC1 the percentage of households in default is observed to be significant both for the model 'As at PR19' and for 'Data to 2019', but not when the time period is extended to include data up to 2022. We would expect households in default to be a relevant driver of bad debt costs, supported by the variable's significance in the shorter time period models. It is possible that the inclusion of the Covid-19 period, and the observed significant changes in bad debt related costs then, is affecting that result; particularly so when, as is the case in these models, no allowance is made year-specific effects
- We observe lower R-squared for the updated dataset models for both RDC1 and RDC2, notably so for the "Data to 2022" model.
- For the RDC2 models the customer transience/migration statistical values are significantly different to those reported by Ofwat at PR19, alongside being insignificant for both data to 2022 and 2019, despite being significant at the 10% level in Ofwat's output.
- Overall, the updated dataset with data to 2022 finds fewer variables to be significant, compared both to Ofwat's PR19 reporting and the data to 2019 model from the same dataset.

Other retail cost models

Table 17 sets out the model results for other retail costs models.

Table 17 Variants of Ofwat's PR19 other retail cost models (ROC)

Model Ref	ROC1			ROC2		
Model variant	As at PR19	Data to 2019	Data to 2022	As at PR19	Data to 2019	Data to 2022
Dependent variable	Ln (Other retail costs per household)					
Explanatory variables						
Percentage of dual service customers	0.002*	0.002** (1.991)	0.002** (2.235)	0.002**	0.004*** (3.018)	0.003*** (3.817)
Percentage of metered connections	0.007***	0.006*** (2.823)	0.001 (0.717)	0.007***	0.006** (2.827)	0.001 (0.726)
Total households connected				-0.039	-0.072 (-1.587)	-0.049 (-1.552)
Overall R-squared	0.13	0.14	0.13	0.15	0.16	0.15
Observations	105	102	153	105	102	153

*Note: For each explanatory variable included in a model, the table reports: value of estimated coefficient; in brackets, the z-statistics; and whether the coefficient is significant at the 1, 5 or 10 per cent significance level, indicated by ***, ** and * respectively.*

Comparing the results set out in the table, we note the following:

- Unlike bad debt models we observe a higher R-squared for the models employing the updated dataset using data to 2019, although the increase is very slight.
- We make a similar observation to the bad debt models; that the model data to 2022 loses significance for variables in comparison to both the model as at PR19 and data to 2019. This may again be due to the unmodelled effects of Covid-19.

Models of total retail costs

Table 18 overleaf sets out the model results for total retail cost models, drawing on the structure outlined at the start of this section.

With regard to the results set out in Table 3 above we note that the values of R-squared are very similar for ‘As at PR19’ and ‘Data to 2019’ across all models, however significantly lower when the model is estimated on data to 2022. As noted for the bad debt and other retails costs, this may be due to the effects of Covid-19.

Table 18 Variants of Ofwat’s PR19 total retail cost models (RTC)

Model Ref	RTC1			RTC2			RTC3		
Model variant	As at PR19	Data to 2019	Data to 2022	As at PR19	Data to 2019	Data to 2022	As at PR19	Data to 2019	Data to 2022
Dependent variable	Ln (Total retail costs per household)								
Explanatory variables									
Total household revenue per customer	0.458***	0.493*** (6.018)	0.505*** (6.757)	0.526***	0.620*** (5.476)	0.595*** (6.540)	0.603***	0.718*** (5.757)	0.619*** (5.589)
Percentage of metered households	0.004	0.001 (0.375)	0.002 (0.470)	0.004	0.003 (0.915)	0.003 (0.905)	0.002	0.001 (0.442)	0.000 (0.122)
Percentage of households in default	0.024	0.014 (0.901)	0.009 (0.682)	0.030**	0.028* (1.950)	0.018 (1.267)			
Total households connected				-0.059*	-0.98** (-2.363)	-0.072** (-2.442)	-0.116**	-0.174** (-2.924)	-0.059* (-1.787)
Income score (unadjusted)							0.059***	0.060*** (3.162)	0.000 (0.008)
Customer transience/migration							0.037**	0.038** (2.020)	0.003 (0.235)
Overall R-squared	0.67	0.66	0.60	0.70	0.70	0.62	0.71	0.70	0.59
Observations	105	102	153	105	102	153	105	102	153

Overall findings

- Across all PR19 models, we observe a significant decrease in both the variables seen as statistically significant and the levels of significance for ‘Data to 2022’.
- Related to the above, on the whole, Ofwat’s model specifications ‘As at PR19’ tend to present slightly higher R-squared scores than models relying on the updated dataset, with the exception of the other retail cost models.

A2 Review of Ofwat's approach to inflation indexation for residential retail price controls

Review of Ofwat's approach to inflation indexation for residential retail price controls

Report for Bristol Water and Wessex Water – 6 April 2023

Reckon LLP

Data analysis | Economic regulation | Competition law

Table of contents

1. Introduction and summary 3
2. Ofwat's explanation of its approach to retail indexation 6
3. Review of Ofwat's reasoning on residential retail indexation 11
4. Interactions with the cost of capital 17
5. A possible way forward for PR24 19

1. Introduction and summary

Background

As part of a project for Bristol Water and Wessex Water concerning residential retail cost assessment, we carried out a review of the approach that Ofwat used at PR19. One specific topic that Bristol Water and Wessex Water were interested in was Ofwat's decision to set the retail revenue allowances in nominal terms, with no CPIH indexation over the five-year price control period.

In its PR24 final methodology Ofwat confirmed that it plans to retain this aspect of its PR19 approach for PR24.

We agreed with Bristol Water and Wessex Water to carry out some further work which considered the indexation (or lack of indexation) of residential retail price controls in more detail.

Ofwat's explanations for not providing inflation indexation

In UK price control regulation, especially when controls are set for a period of five years or so, it is standard practice to set controls on revenues or prices in a way that means that these are automatically adjusted in line with a published index of consumer price inflation (e.g. CPIH and previously RPI).

Historically, Ofwat's price controls for water companies were subject to RPI inflation. When it introduced separate price controls for water companies' residential retail activities at the PR14 review, Ofwat set these revenue controls in nominal terms with no inflation indexation. Ofwat maintained this approach at PR19. Similarly, Ofwat plans no inflation indexation for the residential retail control at PR24.

We have reviewed the explanations given by Ofwat in key documents from its PR14 and PR19 price reviews and from its draft and final methodology for PR24. Ofwat seems to have drawn on four main lines of argument:

- **The link between inflation indexation and asset intensity**. While Ofwat has retained inflation indexation for the wholesale price controls, it associated this decision with the large capital investment programmes for wholesale activities and did not see a similar need for inflation indexation for residential retail activities which it considered to be much less capital intensive.
- **Consistency with the conditions faced by retailers elsewhere in the economy**. Ofwat said that not providing for inflation indexation of the retail control would be in line with the experiences of most retail businesses.
- **The controllability of the input prices faced by water companies**. Ofwat said that, in its view, the inflation risk for water retailers mainly consisted of labour costs and that labour costs are controllable by water companies.
- **Incentives for efficiency in retail operations**. Ofwat said that not allowing inflation indexation would lead to better incentives for water companies to operate efficiently in their retail activities.

Key points from our assessment

Based on the assessment and discussion presented in this report, we draw out the following key points:

- There is no single correct answer to whether there should be some form of inflation indexation – or other adjustment mechanism for unexpected inflation or input price changes – of the retail control. This is a matter of judgement. But this judgement should be made in light of a sound understanding of the relevant arguments and considerations.

- Some of Ofwat's arguments against indexation from PR14 and PR19 do not stand up to scrutiny, specifically the first three arguments listed in the subsection above. Some relevant considerations relating to the impacts on financing costs also seem to have been overlooked.

- We consider that Ofwat was right at PR14 and PR19 to draw on comparisons with the conditions faced by retailers in competitive parts of the UK economy in deciding whether to apply inflation indexation (but we disagree with its interpretation of those conditions).

- There is a reasonable concern that automatic CPIH indexation of retail revenue controls would be overly generous to water retailers compared to the conditions faced by retailers in competitive parts of the UK economy and would lead to unnecessarily high costs to customers in scenarios of unexpectedly high inflation.

- Not allowing for any form of inflation indexation – or other adjustment mechanism for unexpected inflation or input price changes – seems well out of line with the conditions faced by retailers and customers in competitive parts of the UK economy.

- No indexation or other adjustment imposes what seems to be an abnormally large amount of inflation risk on water retailers. And the inflation risk protection that Ofwat is arranging, on behalf of customers, has a cost to customers in terms of the associated financing costs for companies. That high degree of inflation risk protection is not one which consumers usually choose to pay for. Consumers do not typically fix prices for their retail services on a nominal basis for five-year periods.

- Ofwat's decision not to allow CPIH indexation of residential retail controls at PR24 is at odds with its decision to allow CPIH indexation in the calculation of default tariffs for non-residential activities in England. Ofwat does not seem to have provided a good explanation for the differences in its regulatory approach across these two areas.

- At PR19, in explaining its decision not to allow inflation indexation for the retail control, Ofwat emphasised concerns that allowing such indexation would harm water companies' incentives to operate efficiently in their retail activities. This strikes us as a somewhat unusual position in the context of UK price control regulation. We can see arguments to support it, but there are questions of whether this issue is material and whether providing no inflation indexation or other adjustment mechanism is a proportionate response to these concerns.

- Whatever approach is taken on this matter should be properly taken account of in Ofwat's assessment of the retail margin and, perhaps more importantly, the adjustment applied to the appointee WACC to calculate the wholesale WACC.

Intermediate options for price control inflation risk

In light of the points of comparison with retailers in competitive parts of the UK economy and Ofwat's arguments on efficiency incentives, we suggest that some form of compromise approach is an option for PR24 that deserves serious attention.

We identify three different types of approach:

- **Option 1: Inflation risk-sharing mechanism**. Under this approach, there would be a form of CPIH indexation of the residential retail price controls, but in a way that shares the impact of unexpected increases (or decreases) in the CPIH inflation rate between companies and customers.
- **Option 2: Targeted adjustment mechanisms for specific cost categories**. Under this approach, there would be no adjustments to the residential retail price control according to CPIH inflation, but there would be one or more targeted mechanisms for adjustments in respect of specific retail cost categories. For instance, there could be a labour input price adjustment mechanism along similar lines to that used by Ofwat and the CMA for the wholesale controls.
- **Option 3: Adjustment mechanism based on outturn industry-wide retail costs**. Under this approach, there would be no adjustments to the residential retail price control that are based on economy-wide inflation measures or on input price data from outside the water industry. Instead, there would be an adjustment based on outturn unit costs observed across water retailers during the price control period.

We describe these options in more detail and provide a brief review of them against the current approach for residential retail controls.

Structure of this report

The remainder of this report is organised as follows:

- Section 2 provides background information on Ofwat's explanation for its approach of not providing for inflation indexation of the residential retail price controls, drawing on key regulatory publications from the PR14, PR19 and PR24 reviews.
- Section 3 provides a review of Ofwat's stated reasoning for not allowing inflation indexation for the residential retail price controls.
- Section 4 briefly discusses some interactions with the cost of capital.
- Section 5 summarises our assessment emerging from sections 3 and 4, and then suggests some alternative options for PR24 which might make more sense than the current approach.

2. Ofwat's explanation of its approach to retail indexation

In this section, we provide background information on Ofwat's explanation for its approach of not providing for inflation indexation of the residential retail price controls.

We look chronologically at Ofwat's approach since PR14, taking the following in turn:

- Ofwat's approach to indexation when retail controls were introduced at PR14.
- Ofwat's approach to indexation of the residential retail controls at PR19.
- Ofwat's proposed approach to indexation of the residential retail controls at PR24.

This section is purely descriptive, intended to capture the key arguments and explanations provided by Ofwat to justify its approach to retail price control indexation since PR14. We review the arguments and explanations in section 3.

Ofwat's approach indexation when retail controls were introduced at PR14

Ofwat first introduced separate price controls for residential retail activities at the PR14 price review, with effect from April 2015. Prior to this, residential retail activities and wholesale activities were subject to an integrated form of price control, which was subject to RPI inflation indexation during each year of the price control period.

Ofwat decided at a relatively early stage of the PR14 price review process that it would not apply any form of inflation indexation to the residential retail price controls. It said the following in its final methodology in July 2013:[1]

> *"have already confirmed that we will allocate all of the existing RCV to the wholesale business, including for existing retail assets, and have already confirmed that for the less asset intensive retail businesses, RPI indexation is not appropriate."*

When it had previously consulted on its approach, Ofwat explained its position as follows:[2]

> *"Our approach to setting retail controls differs fundamentally from our approach to wholesale controls; this reflects the different nature of the retail and wholesale businesses, and in particular the fact that retail activities use relatively few (and relatively short-lived) assets. That is why a price control based on a regulatory capital value (RCV) and index linked to RPI is not an appropriate form of control for these types of activities."*
>
> *"In the previous controls, we have addressed input price risk by allowing companies to increase prices automatically every year by the change in the Retail Price Index (RPI) with an adjustment for expected efficiency (the K factor). This reflected our view that RPI was a good proxy for the uncertain input price risks faced by a vertically-integrated water*

---

[1] Ofwat (2013) *Setting price controls for 2015-20 – final methodology and expectations for companies' business plans*, page 18.

[2] Ofwat (2013) *Setting price controls for 2015-20 – framework and approach: a consultation*, page 91.

company, taking account of the significant proportion of capital costs and the need to ensure recovery of the costs of long-lived assets in future time periods. But retail services alone require a very different mix of costs compared with wholesale, a much lower proportion of capital costs. They also have assets with much shorter lives and existing retail assets in the RCV have been allocated to the wholesale price control, which will continue to have RPI indexation. So, we do not consider that RPI indexation would be appropriate for the household retail control – and note that this would be in line with the experiences of most retail businesses, across the economy. We expect retailers to manage emerging cost pressures actively by seeking lower costs and optimising their mix of retail inputs, as all retail business in other sectors would. For example, IT systems tend to both fall in cost over time and increase in capability allowing retailers to maintain or improve service at lower cost. We note that some water companies make use of outsourcing and this model is also extensively used to provide retail services in other sectors. We would expect efficient retailers to innovate and actively consider alternative means of providing retail services and so identify scope to reduce costs and (or) improve services. If stakeholders consider that the household retail price control does need to allow for uncontrollable input price pressures, then we would welcome evidence from them about what these costs are and how they meet the three criteria set out in section 5.2.4. If we were to provide a mechanism to make such allowances, then there are a number of approaches we could take that did not involve indexation by RPI and these various approaches are set out in section 6.5."

In the extract above, Ofwat cross-referred to approaches that would not involve RPI indexation but would allow for uncontrollable input price pressures. These are described later in a separate part of its consultation, in the context of non-residential retail controls are as follows:[3]

- Ensuring that the non-household retail control has sufficient net margin to cover the risks of unexpected uncontrollable changes in input prices; and
- A pre-set measure that reflects the future changes in relevant efficient costs of a retailer providing non-household services.

In its PR14 final determinations, Ofwat adopted a very similar line of reasoning for its decision not to allow indexation of the residential retail controls:[4]

"[Unlike] the wholesale price controls, the retail price controls are not automatically indexed to RPI as we do not consider that automatic indexation is appropriate for the less asset and capital-intensive retail businesses. This is more consistent with the arrangements you would expect elsewhere in the economy where retailers do not see prices automatically indexed to RPI. We have required companies to make an evidence-based case for additional revenue for future price pressures in 2014-15 and beyond and have only allowed these changes to household retail upper quartile efficient companies. Furthermore, for those upper quartile companies that have provided sufficient and

---

[3] Ofwat (2013) *Setting price controls for 2015-20 – framework and approach: a consultation*, page 107.

[4] Ofwat (2014) *Setting price controls for 2015-20 Final price control determination notice: policy chapter A5 – household retail costs and revenues*, pages 19-20.

convincing evidence that they are facing input price pressure costs, we have only allowed for a reasonable amount of input cost pressure based on a bottom up assessment of the input cost pressures that they are likely to face in their retail businesses, rather than simply allowing RPI. This is designed to be as consistent as is currently possible with what might happen in a competitive retail market. This has brought further downward pressure on customers' bills."

Ofwat's approach to indexation of the residential retail controls at PR19

At PR19, Ofwat maintained the approach from PR14 of not allowing any inflation indexation of the residential retail price controls.

Interestingly, Ofwat's stated rationale for not allowing inflation indexation seemed to change or evolve between PR14 and PR19. As shown above, at PR14 Ofwat had placed emphasis on the point that water companies' retail activities are less asset intensive than wholesale activities, and made broad references to consistency with arrangements elsewhere in the economy where retailers do not see prices automatically indexed to RPI.

At PR19, Ofwat's stated reasoning gave explicit weight to the view that no indexation would provide better incentives for companies to operate efficiently and – related to this – Ofwat claimed at PR19 that the inflation risk faced by water companies is controllable by water companies.

In its PR19 final methodology, Ofwat highlighted the following arguments:

> *"We consider that not automatically indexing retail controls to inflation provides better incentives for retailers to manage input prices and a more appropriate allocation of risk between customers and companies. We will consider whether any allowance for input inflation needs to be made as part of totex."*[5]

> *"As at PR14 we will not index the retail controls to a measure of general inflation at PR19. We remain of the view that this approach continues to provide appropriate incentives for companies to manage retail input costs. In retail controls the relevant inflation risk is the risk that input prices increase in the short term. There is no RCV in retail controls, so indexation is only relevant to allowed revenue and is not needed to protect the long-term value of the RCV against long-term inflation risk. We remain of the view that inflation risk for water retailers, which mainly consists of labour costs, is controllable by companies. If robust evidence demonstrates that input price pressures present a material cost to companies at PR19, we will assess this as part of our totex allowance, not through indexation."*[6]

---

[5] Ofwat (2017) *Delivering Water 2020: Our final methodology for the 2019 price review*, page 136.
[6] Ofwat (2017) *Delivering Water 2020: Our final methodology for the 2019 price review*, page 154-155.

Elsewhere in its PR19 methodology, Ofwat expanded on some of these points:[7]

> *"We will not index the retail controls to a general measure of inflation. We consider that this approach is most appropriate for the retail controls, and provides appropriate incentives for companies to manage input costs. This is consistent with the incentives for businesses in more competitive markets.*
>
> *We consider that inflation risk in our retail controls is low relative to inflation risk in wholesale controls. In retail, a relatively large proportion of retail costs is attributed to labour costs, and companies should be incentivised to manage the risk of labour cost pressure in the short term. Moreover, the fact that there is no RCV in the retail controls means that indexation is only relevant to the allowed retail revenue. Indexation is not needed to protect the long-term value of the RCV against long-term inflation risk.*
>
> *We will review evidence on forecast input price pressure in retail for the duration of the price control. If appropriate, we will make a cost allowance for inflation as part of totex. This approach ensures companies stay incentivised to manage the risk of input price pressure. We will consider evidence on input price pressure submitted by companies6. We will also consider independent data sources and forecasts, such as data from the Office for National Statistics on wage growth rates. Given that our PR19 approach involves setting an efficient cost allowance for all companies, we intend to apply a common method for determining an inflation allowance for all companies, if we consider that such an allowance is appropriate."*

Ofwat adopted similar reasoning in its PR19 final determinations when it confirmed that it would not apply indexation to residential retail controls and referred back to the PR19 methodology.[8]

Ofwat's proposed approach to indexation of the residential retail controls at PR24

In its PR24 draft methodology, Ofwat said that it planned to retain the approach from PR19 of not applying inflation indexation to the residential retail control:[9]

> *"We propose no automatic indexation of allowed revenue. This means that companies' allowed revenue would not change in line with the general level of inflation within the price control period. Instead, expected input price pressure would be reflected in the revenue limit we set for companies at the outset of the price control. We consider that this approach provides appropriate incentives for retailers to manage input costs given they are best placed to manage them. This is the same approach we took at PR19."*

In its PR24 final methodology, Ofwat reported the following concerns raised by water companies on this issue:[10]

---

[7] Ofwat (2017) *Delivering Water 2020: Our final methodology for the 2019 price review Appendix 11: Securing cost efficiency*, page 22-23.

[8] Ofwat (2019) *PR19 final determinations: Securing cost efficiency technical appendix*, page 129.

[9] Ofwat (2022) *Creating tomorrow, together: consulting on our methodology for PR24*, page 35.

[10] Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24*, page 39.

"[Many] water companies and WaterUK strongly disagreed with our proposal to retain our approach of not automatically indexing allowed retail revenue for inflation. Points raised by these stakeholders included that most retail costs vary with inflation; inflation risks would continue to lie fully with companies; high inflation would risk curtailing service improvements if companies reduce costs; and automatic indexation is a feature in the English business retail market".

Ofwat's PR24 final methodology did not respond in any detail to the concerns it reported, and Ofwat simply explained its decisions for the PR24 methodology as follows:[11]

"Although there were concerns about a lack of automatic retail indexation, any expected input price pressure can be reflected in the revenue limit we set for companies at the outset of the price control instead. This approach provides appropriate incentives for retailers to manage input costs given they are best placed to manage them while still ensuring expected inflationary pressures on input prices can be reflected in companies' allowed revenues. It is appropriate for our PR24 retails controls which (unlike the Retail Exit Code covering the English business retail market) are for a pre-defined duration.

As at PR19, Ofwat seems to have given emphasis to the view that not providing for inflation indexation provides better incentives for companies to act efficiently in their retail activities.

[11] Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24*, pages 39-40.

3. Review of Ofwat's reasoning on residential retail indexation

In this section, we provide a review of Ofwat's reasoning for not allowing inflation indexation for the residential retail price controls. We organise our review around a number of different themes, which relate to different aspects of the case against indexation that Ofwat has raised since PR14, as we have summarised in section 2. We take the following themes in turn:

- The link between inflation indexation and asset intensity.
- Consistency with the conditions faced by retailers elsewhere in the economy.
- The controllability of the input prices faced by water companies.
- Incentives for efficiency in retail operations.

Overall, we find that for the first three points above, Ofwat's stated reasons for not allowing inflation indexation for the residential retail price controls do not stand up well to scrutiny. Furthermore, as discussed further in section 4, Ofwat's reasoning does not seem to take account of relevant considerations relating to the cost of capital. That said, Ofwat's interest in the second point, and in the fourth point, does raise valid questions about the suitability of full CPIH indexation of retail revenue allowances – despite the limitations of Ofwat's explanations of its approach, the case for full CPIH indexation is not a simple open-and-shut one.

The final subsection briefly comments on Ofwat's stated explanation for adopting a different approach for residential retail activities compared to its approach for non-residential retail activities (in England) where it allows inflation indexation for the calculation of default tariffs.

The link between inflation indexation and asset intensity

Ofwat has given weight to the view that inflation indexation was not needed for the retail price controls because retail activities are not asset intensive and because water companies' legacy RCV has been allocated entirely to the wholesale controls.

We would agree with the position that, because of the scale of water companies' wholesale investment programmes, and the size of companies' RCV, there is a good argument for indexation of wholesale controls as a means to ensure investor confidence and – in turn – enable customers to benefit from a lower regulatory cost of capital allowances than would be the case in the absence of indexation.

However, at PR14 Ofwat seemed to treat capital intensity as the primary reason why a regulator might allow inflation indexation of a price control. Ofwat's argument was essentially that, because the specific features or circumstances that called for inflation indexation of wholesale controls did not apply to the retail controls, there was no need for inflation indexation at the retail level. This does not seem to be a good way to think about things. Just because one specific reason for inflation indexation at the wholesale level does not apply at the retail level, it does not follow that there are no grounds for inflation indexation at the retail level.

For instance, some potential arguments in favour of inflation indexation are as follows:

- It is not normal for companies (retail or otherwise) to operate under financial constraints which prevent them from increasing prices at any point over a five-year period even in the presence of substantial levels of inflation across the economy.

- Not taking steps to protect – to some degree at least – water companies from the financial risk associated with potential scenarios of unexpectedly high inflation may feed through to higher charges to customers (e.g. through the allowed cost of capital) and may force customers to pay for excessive levels of inflation risk protection.

Neither of these points relate closely to asset intensity. Companies with few assets can face substantial inflation risk if locked into fixed nominal pricing across five or so years (e.g. in terms of wage rates).

From its published documents, it does not seem that Ofwat properly considered the case for and against indexation, having adopted a narrow view of the rationale for indexation that was tied to capital intensity.

Consistency with the conditions faced by retailers elsewhere in the economy

Ofwat said at PR14 that not providing for inflation indexation of the retail control "*would be in line with the experiences of most retail businesses, across the economy*".

This is true in a very literal sense. Nonetheless, the effect of Ofwat's approach to indexation is to impose conditions (and risks) on water company retail businesses that are firmly out of line with the experiences of most retail businesses.

It is probably true that most retail businesses across the economy do not enjoy the benefit of automatic inflation indexation of the prices that they can charge their customers each year.

But nor do they operate under the fundamental constraints arising from the system of price control regulation applied to water companies, which tightly restrict the level of revenue that water companies can collect from customers each year. The question of indexation must be viewed within the context of price control regulation.

While retailers elsewhere in the economy may not enjoy automatic inflation indexation of their prices, they have far greater flexibility on their pricing than a regulated company subject to fixed term price controls. It seems highly unlikely that, in a competitive market, a retailer that is operating in conditions in which there is 10% retail price inflation across the economy, and substantial increases in nominal wage rates, would be unable to increase its prices over a five-year period. While retailers would not necessarily pass on the full amount of economy-wide inflation immediately to customers, we would expect at least some of this to be passed on to customers over time.

On this basis, Ofwat's position that no inflation indexation "would be in line with the experiences of most retail businesses, across the economy" does not seem valid and we do not consider that it would not stand up to review (e.g. by the CMA).

That said, we do think that comparisons with the conditions that might be faced by retailers elsewhere in the economy provide a helpful perspective on the question of whether to apply inflation indexation (or some other form of adjustment mechanism) to retail price controls.

While a detailed analysis of inflation experienced by retailers in competitive sectors is beyond the scope of this report, our view is that unexpected high inflation in competitive markets is likely to have a number of different impacts, rather than being simply borne in full by either retailers or customers. For instance, there are several ways in which customer prices might not bear the full extent of unexpected economy-wide inflation, at least in the short term, such as:

- differences in the inflation rate for a retailer's inputs relative to economy-wide inflation (e.g. if CPIH is driven in large part by energy prices and a retailer's inputs and supply chain reflects energy costs to a lesser degree);
- reductions in inflation-adjusted wages and salaries as labour does not necessarily have bargaining power in a context of uncertainty about the nature and duration of inflation to insist on increases that maintain wage levels in real terms (and, similarly, reductions in inflation-adjusted payments to the retailer's suppliers);
- reductions in the inflation-adjusted profit of retailers due to competitive constraints between them limiting price increases, especially in a context of uncertainty about the nature and duration of the high inflation situation and about how competitors will respond to this; and
- inflation acting as a spur to efficiency improvements, and opportunities for input substitution.

Nonetheless, we cannot see a good economic basis for expecting that, in a competitive market, customer prices will not adjust at all in response to changes in economy-wide inflation.

On this basis, a more credible comparison of water companies' retail businesses with retailers in the wider economy would be as follows:

- Automatic inflation indexation of the five-year retail control would provide more favourable and lower-risk conditions than are typically faced by retailers elsewhere in the economy.
- No inflation indexation of a five-year retail control (or similar adjustment mechanisms) would provide harsher and riskier conditions than are typically faced by retailers elsewhere in the economy.

There is likely to be relevant evidence on where competitive parts of the UK economy fall between these extremes arising from the current period of very high inflation (e.g. profiles of wage and earnings data versus CPIH or profiles of sectoral output price metrics versus CPIH).

The controllability of the input prices faced by water companies

Some aspects of Ofwat's reasoning since PR14 seem to relate to the view that the input prices faced by water companies are controllable.

As highlighted in section 2, Ofwat said the following in its PR19 final methodology:[12]

> *"We remain of the view that inflation risk for water retailers, which mainly consists of labour costs, is controllable by companies."*

[12] Ofwat (2017) *Delivering Water 2020: Our final methodology for the 2019 price review*, page 154-155.

This statement seems relevant to the reasonableness of Ofwat's decision not to apply inflation indexation or any other adjustment mechanism.

It is clearly true that companies have a significant degree of influence over the wages and salaries that they pay to their employees.

But it is equally true that there is such a phenomenon as economy-wide inflation which reflects, for example, the effects of UK consumer and business sentiment, the actions of the Bank of England, and events in other countries. This phenomenon as economy-wide inflation has a substantial impact on the labour costs faced by efficient companies.

Ofwat's statement that labour costs are controllable by companies is too simplistic and risks overstating the extent to which efficient water retailers are able to control labour costs over a five-year price control period. In turn, if Ofwat overestimates the extent to which efficient water retailers are able to control labour costs, it risks making misguided decisions on whether or not to apply inflation indexation to the residential retail control.

In addition, Ofwat's stated view that inflation risk for water retailers mainly consists of labour costs does not seem valid. Around 40% of residential retail costs are costs related to doubtful debt.[13] Both economic logic, and evidence from econometric models of bad debt costs, suggest that there will be a relationship of approximately 1:1 between the value of total billed revenue and bad debt related costs. Given the CPIH indexation that Ofwat applies to wholesale revenues, and the relative scale of wholesale revenues to retail revenues, total billed revenues will vary over the price control period in response to changes in CPIH versus forecast CPIH (and for other factors, such as revenue adjustments for in-period ODIs). These changes seem outside of the control of an efficient water retailer yet are likely to have a direct effect on an efficient water retailer's bad debt related costs.

Incentives for efficiency in retail operations

As shown in section 2, at PR19 Ofwat gave more emphasis in its explanation for not allowing inflation indexation of the retail controls to the view that this would lead to better incentives for water companies to operate efficiently in their retail activities.

This line of argument is a little surprising as, in our experience, the established practice in UK price control regulation does not treat inflation indexation as creating a significant risk for the efficiency of regulated companies.

However, there are some arguments that might be used to elaborate on the concerns expressed by Ofwat about inflation indexation posing risks to water companies' efficiency (though Ofwat did not express these points):

- It might be argued that the tougher the price control is on water companies, the stronger are the financial incentives on them to be efficient, and not allowing indexation leads to tougher controls in scenarios of unexpectedly high inflation.

---

[13] Covering debt management costs and bad debt costs.

- It might be argued that, in periods of relatively high and unexpected economy-wide inflation, water companies are less likely to get a good deal from their suppliers and to limit nominal rises in the wages and salaries they pay, if the retail business benefits from inflation-linked revenues. For instance, it may be more difficult to negotiate below-inflation increases in payments to suppliers and employees to the retail business if the latter know that company revenues for retail activities are automatically increased in line with inflation. This line of argument is particularly relevant if there is evidence that, in practice, suppliers and employees of retail businesses elsewhere in the economy do not receive increases in remuneration that matches inflation in high-inflation scenarios.

The first argument would imply regulatory arrangements that are intended to create stronger incentives on companies in scenarios of unexpectedly high inflation than in more typical scenarios around inflation, which would then raise questions about the adequacy of incentives in normal conditions. This does not seem to be a good justification for not allowing inflation indexation.

The second argument seems more plausible, though there are questions about the significance of any such effects. Furthermore, the hypothesis that a company's prices being indexed to an inflation measure such as CPIH or RPI poses a material threat to its own efficiency seems somewhat at odds with the observation that such indexation is adopted voluntarily by companies operating in competitive parts of the economy who operate under multi-year contracts with their customers.

In any event, there is then a question of whether allowing no inflation indexation (or similar adjustment mechanism) is a proportionate response to the hypothetical risks to efficiency incentives. This is especially so in a context where wider aspects of retail price controls, such as the emphasis on benchmarking costs and performance, should provide quite strong incentives for companies to operate efficiently.

Ofwat's distinction between residential and non-residential retail price controls

In its decision on the review of the Retail Exit Code for non-residential retail activities in December 2022, Ofwat confirmed that it would retain its approach of allowing for CPIH-indexation of the retail cost component and meter read component for the default tariff calculations for Group One customers.[14]

In its PR24 final methodology, Ofwat mentioned that some of the objections to its proposals not to allow inflation indexation of the residential retail price controls referred to the fact that allowing for inflation indexation is a feature of Ofwat's approach to regulation of the English business retail market.

Ofwat seemed to dismiss the idea that there was an inconsistency in its regulatory approach across residential and non-residential retail activities, explaining stating that its approach for the residential retail controls "… *is appropriate for our PR24 retail controls which (unlike the Retail Exit Code covering the English business retail market) are for a pre-defined duration*".[15]

---

14 Ofwat (2022) *Business retail market 2021-22 review of the Retail Exit Code - Decision Main document*, page 24.

15 Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24*, pages 39-40.

We consider this argument to be quite weak. This is especially so in a context where Ofwat said that it plans to review the price and non-price protections in the Retail Exit Code again in 3-5 years' time.[16] This means that, on Ofwat's own plans, the non-residential retail controls which feature inflation indexation are intended to last for no longer than the PR24 residential retail controls (which do not feature indexation) and that the non-residential retail controls could be updated or replaced much sooner than the residential retail controls.

While not directly stated by Ofwat, a possible reason why Ofwat wanted to retain indexation for the non-residential default tariffs is that, given the policy of developing competition in this area, Ofwat attached greater importance to the cost allowances underpinning price controls being reflective of changes over time in efficient costs than in the case of the residential retail controls.

16 Ofwat (2022) *Business retail market 2021-22 review of the Retail Exit Code - Decision Main document*, page 3.

4. Interactions with the cost of capital

Our view is that a proper assessment of whether or not to allow for inflation indexation should consider the interactions between inflation indexation and the cost of capital for a notional efficient retail company (or notional efficient water company).

All else equal, we would expect a notional efficient retailer to face a higher cost of capital if its price control is fixed in nominal terms compared to the case where it is indexed to CPIH. We see no reason to think that risks relating to the impact on a water company's retail costs from uncertainty around economy-wide inflation are fully diversifiable from the perspective of investors. Furthermore, a notional efficient retail company that is exposed to this risk may need to operate at lower gearing to provide sufficient equity buffer to accommodate high-inflation scenarios than a company benefitting from CPIH indexation of revenues.

The interactions between the approach to inflation indexation and the cost of capital have two main implications for the setting of residential retail price controls:

- First, if the potential for interactions with the cost of capital are overlooked, the regulatory judgement on whether or not to apply inflation indexation to the retail price controls seems unlikely to be being taken on a sound basis.
- Second, even if there were to be good reasons not to allow inflation indexation of the retail control there is a question of whether retail allowances (or companies' total allowances across retail and wholesale) provide sufficient compensation for the risks (or the costs of risk mitigation) arising from a price control structure that involves revenue allowances fixed for five years irrespective of the level of inflation in the UK economy.

On the first point, there is an argument that no inflation indexation exposes customers to disproportionate financing costs. This argument is particularly valid in a context where no indexation or other adjustment imposes what seems to be an abnormally large amount of inflation risk on water retailers, compared to that experienced in competitive sectors.

Put differently, the inflation risk protection that Ofwat is arranging, on behalf of customers, has a cost to customers in terms of the associated financing costs for companies. That high degree of inflation risk protection is not one which consumers usually choose to pay for. Consumers do not typically fix prices for their retail services on a nominal basis for five-year periods.

On the second point, we now turn to consider the extent to which the financing costs associated with no protection against inflation may already be funded via the appointee cost of capital allowance (at least if a similar approach to PR19 is used).

In the context of the financing costs for the residential retail business, both Ofwat and the CMA adopted an approach which broke these financing costs down into three categories: (a) cost of financing fixed assets; (b) required revenue for return on working capital;[17] and (c) required return for systematic risk arising from residential retail activities.[18]

Both the CMA and Ofwat took the view that required return for systematic risk borne by the residential retail business was already funded by the appointee WACC and – to avoid double counting – made a deduction from the appointee WACC based on an estimate of the part of the notional 1% margin for residential retail financing costs that was attributable to funding the financing costs from systematic risk arising from residential retail activities.

In this context, provided that the market data on equity beta that is used for the CAPM analysis feeding into WACC is for a period since PR14 and provided that this data covers listed water companies only, it could be argued that the cost of equity derived from CAPM already reflects the financing costs of residential retail price control arrangements that do not provide companies with any inflation indexation of revenue allowances.

However, there are some reasons why this perspective might be too narrow and inappropriate:

- **Changes over time to perceptions of inflation risk**. The financing costs for systematic risk relating to economy-wide inflation may be higher in the 2025-30 period than in the historical data used for equity beta because, following a recent period of exceptional and unexpected inflation, investors may attach more weight to that risk going forwards than over the previous low-inflation environment.
- **Asymmetric risk around economy-wide inflation**. The risk structure for inflation risk may be asymmetric, in the sense that scenarios of unexpectedly very high inflation may be significantly more likely than scenarios of unexpectedly very low inflation. For instance, using the RPI inflation measure, which is directly available over a longer period, since 1948 there have been a number of episodes, and seven years in total, of RPI annual inflation exceeding 10%, and only a single year in which the annual change in RPI was negative: at just -0.5% in 2009.[19] And from a policy perspective, there seems to be a tendency to treat deflation as a greater economic threat than inflation (e.g. as was manifest in the policy and media discussions during the global financial crisis). As recognised by the CMA, "*the taking on of asymmetric risk cannot in principle be rewarded by applying a WACC determined on CAPM principles, as the CAPM assumes that all risks are symmetrical*".[20]

It is not the role of this report to examine these issues in any detail, but it is relevant to keep them in mind when considering the types of approach that might be suitable for PR24.

---

[17] In contrast to Ofwat, the CMA took the position that an efficient retailer had zero financing costs in respect of component.

[18] CMA (2021) *Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations: Final report*, pages 1016 to 1029.

[19] Reckon analysis of ONS data series CZBH.

[20] CMA (2017) *SONI Limited v Northern Ireland Authority for Utility Regulation: Final determination*, footnote 776

5. A possible way forward for PR24

This section summarises our assessment emerging from sections 3 and 4, and suggests some alternative options for PR24 which might make more sense than the current approach. It takes the following in turn:

- Key points from our assessment.
- Intermediate options for price control inflation risk.
- Discussion of intermediate options.

Key points from our assessment

Based on the discussion in sections 3 and 4, we draw out the following key points:

- There is no single correct answer to whether there should be some form of inflation indexation – or other adjustment mechanism for unexpected inflation or input price changes – of the retail control. This is a matter of judgement. But this judgement should be made in light of a sound understanding of the relevant arguments and considerations.
- Some of Ofwat's arguments against indexation from PR14 and PR19 do not stand up to scrutiny, and some relevant considerations (e.g. impacts on financing costs) seem to have been overlooked.
- We consider that Ofwat was right at PR14 and PR19 to draw on comparisons with the conditions faced by retailers in competitive parts of the UK economy in deciding whether to apply inflation indexation (but we disagree with its interpretation of those conditions).
- There is a reasonable concern that automatic CPIH indexation of retail revenue controls would be overly generous to water retailers compared to the conditions faced by retailers in competitive parts of the UK economy, and would lead to unnecessarily high costs to customers in scenarios of unexpectedly high inflation.
- Not allowing any form of inflation indexation – or other adjustment mechanism for unexpected inflation or input price changes – seems well out of line with the conditions faced by retailers in competitive parts of the UK economy.
- No indexation or other adjustment imposes what seems to be an abnormally large amount of inflation risk on water retailers. And the inflation risk protection that Ofwat is arranging, on behalf of customers, has a cost to customers in terms of the associated financing costs for companies. That high degree of inflation risk protection is not one which consumers usually choose to pay for. Consumers do not typically fix prices for their retail services on a nominal basis for five-year periods.
- Ofwat's decision not to allow CPIH indexation of residential retail controls at PR24 is at odds with its decision to allow CPIH indexation in the calculation of default tariffs for non-residential activities in England. Ofwat does not seem to have provided a good explanation for the differences in its regulatory approach across these two areas.

- At PR19 in explaining its decision not to allow inflation indexation for the retail control, Ofwat emphasised concerns that this would harm water companies' incentives to operate efficiently in their retail activities. This strikes us as a somewhat unusual position in the context of UK price control regulation. We can see arguments to support it, but there are questions of whether this issue is material and whether providing no inflation indexation or other adjustment mechanism is a proportionate response to these concerns.
- Whatever approach is taken on this matter should be properly taken into account in Ofwat's assessment of the retail margin and, perhaps more importantly, the adjustment applied to the appointee WACC to calculate the wholesale WACC.

Intermediate options for price control inflation risk

Looking across the points of comparison with retailers in competitive parts of the UK economy and Ofwat's arguments on efficiency incentives, we suggest that some form of compromise approach is an option for PR24 that deserves serious attention. The aim would be to allow some form of adjustment mechanism for unexpected inflation or input price changes, but in a way that is better aligned with the conditions faced by retailers in competitive parts of the economy than the standard CPIH indexation arrangement seen in UK price control regulation.

In this subsection we briefly outline some alternative options which can be seen to lie on an intermediate position between the extremes of (a) no inflation indexation or other adjustment arrangements and (b) full automatic CPIH indexation of residential retail revenue allowances.

These intermediate options are as follows:

- **Option 1: Inflation risk-sharing mechanism**. Under this approach, there would be a form of CPIH indexation of the residential retail price controls, but in a way that shares the impact of unexpected increases (or decreases) in the CPIH inflation rate between companies and customers. Specifically, the mechanism would involve adjustments, applied in each year of the price control period, for differences between outturn CPIH and the corresponding forecast of CPIH that Ofwat used when it made its final determinations. Different options could be considered around calibration (e.g. 75% or 50% of unexpected variations in inflation passed through to customers). There may be opportunities to draw on evidence on how risks around unexpected levels of inflation are shared between companies and customers in competitive parts of the economy, perhaps drawing on recent experience of high inflation in the UK.
- **Option 2: Targeted adjustment mechanisms for specific cost categories**. Under this approach, there would be no adjustments to the residential retail price control according to CPIH inflation, but there would be one or more targeted mechanisms for adjustments in respect of specific retail cost categories. For instance, there could be a labour input price adjustment mechanism along similar lines to that used by Ofwat and the CMA for the wholesale controls, with calculations based on differences between an outturn wage inflation measure in nominal terms and an upfront regulatory forecast of the corresponding wage inflation measure. There could also be a role for a mechanism to proxy for the impact on an efficient retailer's bad debt costs arising from the CPIH indexation of wholesale price revenue allowances.

- **Option 3: Adjustment mechanism based on outturn industry-wide retail costs**. Under this approach, there would be no adjustments to the residential retail price control that are based on economy-wide inflation measures or on input price data from outside the water industry. Instead, there would be an adjustment based on outturn unit costs observed across water retailers during the price control period. The adjustment to revenue allowances could be based on differences between (a) the average residential retail cost per property (nominal terms) across water companies reported during price control period and (b) the average residential retail cost per property assumed when setting the price control (nominal terms) – or some variant of this comparison. This could be applied as an annual within-period adjustment when data becomes available or as an end-of-period adjustment.

We make some brief comments on the benefits and drawbacks of these options below. A more detailed and systematic assessment is outside the scope of this report.

It is also relevant to recognise that option 3 represents a form of uncertainty mechanism that protects customers and companies against a wider set of uncertainties than those relating to inflation (e.g. the adjustment would tend to capture differences in industry-wide productivity improvements between what is achieved in practice and what was assumed when setting the price control). This broader scope might be seen to have both advantages and disadvantages. Reckon has considered this type of uncertainty mechanism in greater detail in the context of wholesale price controls, as part of previous work which is publicly available.[21]

For each of the three options above, some care is needed to ensure that there is internal consistency between the adjustment mechanism and the way that allowances are set. Our initial view is that, under each of the options, it would make sense for Ofwat to determine retail cost allowances in nominal terms at the price review, and for adjustments to be based on differences between outturn metrics in nominal terms and nominal forecasts of those metrics determined at the price review. This is slightly different to the conventional approach to price control inflation indexation, where allowances are determined in real terms and adjustments are based on outturn CPIH without deducting a forecast of CPIH.

Discussion of intermediate options

At one level, different regulatory approaches to inflation indexation or other adjustments for inflation and input prices are feasible and, from the perspective of companies, what should matter most is that there is a reasonable degree of consistency between the risk exposure that companies face under the chosen and the price control remuneration for the cost of capital (including financing costs associated with retail activities).

Whichever approach is taken – whether any of the above options, full CPIH indexation or the PR19 approach, attention should be given to the interaction with the cost of capital. Options 1, 2 and 3 would tend to reduce companies' risk exposure relative to the current approach and relative to what might be reflected in appointee-level equity beta estimates from historical market evidence. However, as discussed in section 2, there are possible arguments that, under the current approach,

[21] See section 5.6 of Reckon (2022) *The opportunities for a more coherent regulatory approach for Ofwat's funding of base expenditure and enhancements: final report*.

the appointee-level equity beta estimates from historical market evidence would not fully remunerate efficient companies for the financing costs associated with inflation risk exposure. Risks may be seen as higher in light of the recent inflationary experience in the UK and inflation risk may be asymmetric and not remunerated by a cost of equity allowance based on CAPM.

We suspect that it would be difficult to precisely estimate the impact of different options (including the current approach and full CPIH indexation) on the cost of capital faced by water companies which customers need to fund. This may make it difficult, in practice, to make a decision on the choice of approach by comparing the benefits to customers from more stable and predictable retail price controls against the downsides to customers (e.g. slightly higher cost of capital). In any event, since most of the customer bill is made up of wholesale costs, it is questionable whether it is meaningful to envisage customer benefits from fixed retail revenue allowances when wholesale revenue allowances are indexed to CPIH.

In this context, and guided by proportionality, we feel that comparisons with the degree of inflation risk protection in competitive parts of the UK provide a useful reference point in determining an approach for the water retail controls. We think that Ofwat was right to draw attention to such comparisons when explaining its approach at PR19.

In addition to this, we consider that two other important considerations for the choice of approach are the interactions with efficiency incentives and the administrative burden of the approach.

On this basis, we provide some high-level comments on the three options, and the current approach, in the table below. For the purposes of this report, our main interest is how these types of options compare against the current approach, rather than looking in detail at differences between them.

Table 1 High-level comparison of options against relevant considerations

Consideration	Comments
Consistency with conditions faced by retailers in competitive markets	All three options above seem more consistent with conditions faced in competitive retail markets than both (i) the current approach of no inflation indexation or other adjustment mechanism and (ii) the type of full CPIH indexation applied at the wholesale level. Of the three, option 3 is arguably closest to yielding the type of conditions faced by retailers in competitive markets because it effectively uses cross-company benchmarking to determine adjustments.
Impact on efficiency incentives	All three options provide a way to allow some form of adjustment in light of unexpected changes in inflation or input prices while mitigating the concern raised by Ofwat at PR19 about the impacts of full automatic CPIH inflation on retail cost efficiency. They seem to offer a more proportionate response to this concern. There might be concerns raised about efficiency incentives under option 3, due to theories around potential collusion amongst companies, but it is questionable whether these concerns are valid where there are 15 or so independent companies affecting the average cost benchmark.
Simplicity of the approach and administrative burden	All approaches seem reasonably simple within the context of Ofwat's system of economic regulation. Option 1 is the simplest of the three. Option 3 seems likely to be the most complicated to implement well.

A3 Review of Ofwat's use of water companies' cost forecasts in its determination of allowances for residential retail costs

Review of Ofwat's use of water companies' cost forecasts in its determination of allowances for residential retail costs

Report for Bristol Water
and Wessex Water – 6 April 2023

Reckon LLP
Data analysis | Economic regulation | Competition law

Table of contents

1: Introduction and summary....3
2: Context for Ofwat’s use of companies’ retail cost forecasts....7
3: Comparisons of cost forecasts at PR19 and outturn costs....12
4: Review of the case for using forecast cost data at PR24....17

1: Introduction and summary

Background

As part of a project for Bristol Water and Wessex Water concerning residential retail cost assessment, we carried out a review of the approach that Ofwat used at PR19. One specific topic that Bristol Water and Wessex Water wanted us to look at in greater detail concerned Ofwat's use of water companies' business plan cost forecasts as part of its calculation of price control allowances for residential retail activities at PR19. This report presents our review of this issue. Our focus is on residential retail cost assessment, but some issues may have wider implications.

Brief overview of Ofwat's use of business plan forecasts at PR19

At PR19, Ofwat did not draw on water companies' forecasts of residential retail costs to estimate its econometric benchmarking models; those models were estimated using historical data only. But Ofwat did use an adjustment based on a form of cross-company benchmarking of companies' cost forecasts in moving from the results from those models to the cost allowances set for each company.

Ofwat's determination of residential retail cost allowances involved applying a reduction of 15.4% to modelled costs (derived from econometric models applied to historical data) in projecting its cost benchmarks over the 2020-25 period. This reduction was calculated by giving a 50% weight to an upper quartile efficiency adjustment based on companies' cost forecasts (this element implied a 20.6% reduction from modelled costs) and a 50% weight to an upper quartile efficiency adjustment based on historical costs (which implied a 10.2% reduction).

The efficiency adjustment based on cost forecasts (20.6%) was derived from comparisons, across companies, of the ratio of each company's business plan forecasts over the 2020-25 period (in nominal terms) against Ofwat's modelled costs for that company over the same period. The company with the smallest ratio of forecast costs to modelled costs was deemed to have the most efficient cost forecast and other companies were ranked accordingly. Ofwat took the corresponding ratio for the upper quartile company as the basis for its upper quartile efficiency adjustment based on companies' cost forecasts.

The principle of using business plan forecasts in projecting cost benchmarks

Under an approach to cost assessment based primarily on cross-company benchmarking of historical costs, it is necessary to consider how to form projections of the cost benchmarks for each company over the forthcoming price control period. Whether implicitly or explicitly, these projections should take account of changes over time between the projection period and the historical period used for the modelling, in relation to the effects on efficient costs of factors such as ongoing productivity (or efficiency) improvements over time and changes to input prices.

This can be done without using company forecasts at all, for example as part of conventional regulatory assessments relating to ongoing productivity growth (or "frontier shift" efficiency) and real price effects (RPEs). These assessments tend to place substantial weight on historical evidence and on evidence from quite different sectors and economic activities (e.g. regulatory practice often draws on productivity data on the UK as a whole or for specific broad UK sectors for these purposes).

One of the main benefits that we see in using business plan forecasts, in the way that Ofwat did at PR19, is that it provides an alternative method for projecting cost benchmarks over a forthcoming price control period in a way that uses the views of water companies in respect of their residential retail activities, and which can be more forward-looking.

That approach is vulnerable to risks around the quality and reliability of the cost forecasts that companies submit as part of their business plans. Companies' forecasts may be influenced or distorted by a range of financial, procedural and reputational incentives. Indeed, there is a question of whether the forecasts provided by water companies as part of their business plans can reasonably be interpreted as forecasts of their efficient costs – rather than simply as numbers created in response to the incentives created by the price review process.

Evidence from the 2020-25 period so far indicates that, on average, companies have over-spent relative to their business plans forecasts for residential retail costs, with particularly high levels of over-spend for those companies for which Ofwat gave the best grade in its IAP assessment of their business plan forecasts for residential retail costs.

Despite valid concerns about the accuracy of companies' forecasts, we do not see a basis for viewing the use of company forecasts as unreasonable as a matter of principle, at least for the type of adjustment used for residential retail at PR19. All the methods we see for projecting cost benchmarks over a forthcoming price control period have problems. Instead, what seems more important is *(a)* the decision on how exactly forecasts are to be used and *(b)* the steps Ofwat takes to guard against the risks that the allowances it sets for all companies are compromised by certain companies' forecasts being too high or too low.

Mitigating risks relating to the accuracy of business plan forecasts

In a context where Ofwat has reasonable grounds to be concerned that companies' business plan forecasts of retail costs may be unduly high, or insufficiently ambitious, it seems reasonable for Ofwat to adopt an approach that only draws directly on these forecasts if they act to reduce allowances derived from historical data and other sources of evidence and does not use them if doing so would increase allowances. While this might be viewed as cherry-picking, it does not seem unreasonable in the circumstances.

At the same time, we think that it would be unreasonable for Ofwat to act in a way that is blind to the risks that some water companies have submitted unduly low forecasts of retail costs, especially when it is using some companies' forecasts to calculate or adjust retail cost allowances across the industry. This is especially so given the incentives created by the business plan assessment process.

To help mitigate the risk of using company forecasts that are too low, we think that Ofwat needs to pay attention to the credibility of companies' forecasts before these are allowed to influence the cost allowances it sets for other companies. We do not consider it sufficient to point to the assurance requirements for companies' business plans as a guarantee of the credibility and quality of forecasts.

In its PR24 final methodology, Ofwat said that it would adapt its approach to business plan assessment to include a new minimum expectation in its quality assessment for each company to provide sufficient and convincing evidence that it can credibly deliver the proposals in its plan. This should reduce some of the concerns above: we do not see how Ofwat could set price control allowances on the basis of cost forecasts it has found not to have met this minimum requirement.

Even so, it can be difficult and time-consuming to assess the credibility of a company's business plan cost forecasts, and whether it is supported by the evidence provided, and this will require regulatory judgement. There are open questions at this stage as to how effective Ofwat's assessment in this area will be.

In addition to looking at the credibility of forecasts, the concerns about the accuracy of cost forecasts cast doubt on an approach to cost assessment which would give 100% weight to calculations reliant on companies' cost forecasts when setting allowances. Instead, allowances could be set using an average of cost projections involving companies' cost forecasts and one or more approaches that are not reliant on such forecasts. At PR19 Ofwat's approach had the effect of giving 50% weight to cost benchmarks that were adjusted using company forecasts. There may be a case for a lower weighting at PR24, depending on the nature and quality of evidence provided to support them, and on the extent to which Ofwat reviews that evidence.

Upper quartile adjustments

The risks from using unduly low forecasts are exacerbated if Ofwat uses those from a company that is identified as upper quartile in terms of the efficiency of its cost forecasts. On the balance of probabilities, the cost forecasts that Ofwat identifies as being upper quartile (or better) in terms of their perceived cost efficiency are likely to reflect a greater degree of under-estimation of costs than applies on average across the forecasts submitted by all companies. This under-estimation may reflect inadvertent forecasting errors as well as the effects of price review incentives.

Based on two years of outturn data for the 2020-2025 period so far, we found that the average over-spend for the companies ranked by Ofwat at PR19 as being at or above the upper quartile efficiency in terms of their business plan forecasts was 25%, compared to 9% for the other companies.

There may be a way for Ofwat to mitigate, to some degree, this specific concern about the greater forecasting inaccuracy from an upper quartile forecast. Ofwat could first calculate a more conventional upper quartile efficiency adjustment based on historical costs. It could then calculate a further adjustment for dynamic effects (e.g. ongoing productivity improvements and changes to nominal input prices and potentially other factors) calculated as the *average* or *median* across companies of each company's ratio of (a) forecast costs in nominal terms to (b) modelled costs (without applying any efficiency adjustment to these modelled costs). To make projections for the 2025-30 period, both the historical upper quartile adjustment and the further adjustment for dynamic effects would be applied to modelled costs. This approach would enable Ofwat to retain the use of an upper quartile efficiency benchmark (if appropriate), while potentially improving the way that company forecasts are used to bring a forward-looking perspective into the projection of cost benchmarks.

We do not consider in this report broader aspects of the practice of using upper quartile efficiency benchmarks (or similar) to apply adjustments to modelled costs.

Structure of this report

The remainder of this report is structured as follows:

- Section 2 provides some background and context on regulatory practice which is relevant to the use of companies' business plan cost forecasts in setting allowances for water companies' residential retail costs. It provides information on Ofwat's approach for residential retail costs at PR19, Ofwat's use of company forecasts for other aspects of its cost assessment, and some information on how Ofgem has used company cost forecasts in a different way when setting price controls.
- Section 3 presents comparisons of water companies' PR19 forecasts for residential retail costs against outturn costs for the first two years of the price control period.
- Section 4 considers and discusses a number of issues which relate to the case for Ofwat using water companies' forecasts of residential retail costs in settling retail price control allowances at PR24.

2: Context for Ofwat's use of companies' retail cost forecasts

This section provides some background and context on regulatory practice which is relevant to the use of companies' business plan cost forecasts in setting allowances for water companies' residential retail costs. It takes the following in turn:

- Ofwat's use of companies' residential retail cost forecasts at PR19.
- Ofwat's use of business plan forecasts to inform wholesale catch-up at PR19.
- The CMA's approach to the notional company catch-up adjustment at PR19.
- Use of cost forecasts in Ofwat's cost assessment for wholesale enhancements.
- Position on the use of company cost forecasts in Ofwat's final methodology for PR24.
- Ofgem's use of companies' cost forecasts as inputs to econometric models.

Ofwat's use of companies' residential retail cost forecasts at PR19

At PR19, Ofwat did not include the forecasts of residential retail costs as inputs to its econometric benchmarking models; these models were estimated using historical data only. But Ofwat did use an adjustment based on a form of cross-company benchmarking of companies' retail cost forecasts in moving from the results from these models to the retail cost allowances for each company for the 2020-25 period.

In effect, Ofwat's determination of residential retail cost allowances involved applying an adjustment to modelled costs (derived from historical data) which was calculated by giving a 50% weight to an upper quartile efficiency adjustment based on cost forecasts and a 50% weight to an upper quartile efficiency adjustment based on historical costs.

The upper quartile efficiency adjustment based on cost forecasts was calculated as follows. For each company Ofwat calculated the ratio of the company's business plan forecasts for residential retail costs expressed in nominal terms (aggregated over the five-year price control period) and Ofwat's modelled costs for that company over the same period (taking account of forecasts explanatory variables). The company with the smallest ratio of forecast costs to modelled costs was deemed the most efficient and other companies ranked accordingly. Ofwat selected the upper quartile company as the most efficient company.

On this basis, Ofwat calculated a reduction of 20.6%, reflecting the upper quartile company's business plan forecast costs being 20.6% lower in nominal terms than Ofwat's modelled costs for that company (modelled costs were calculated in 2017/18 prices).

Ofwat then combined this figure of 20.6% with an upper quartile adjustment based on historical costs, rather than forecast costs, which implied a reduction on modelled costs of 10.2%. Drawing on these two elements, Ofwat applied a combined efficiency adjustment of 15.4% (calculated as the average of 20.6% and 10.2%) to modelled costs for each company for each year in the 2020-25 period. This meant that, in setting residential retail cost allowances at PR19, Ofwat effectively gave 50% weight to an adjustment calculated using companies' business plan cost forecasts.

Ofwat's use of business plan forecasts to inform wholesale catch-up at PR19

In setting wholesale controls at PR19, Ofwat applied what seems to be a more conventional notional efficient company catch-up adjustment, calculated using comparisons of each company's modelled costs relative to its outturn cost over the previous five years.

However, while this adjustment was ostensibly an adjustment based on evidence on historical performance, in practice Ofwat took account of forward-looking evidence too – specifically water company business plans – in deciding what company to use as the notional efficient company for the purposes of calculating the adjustments for wholesale water and wholesale wastewater at PR19. This, in turn, affected the scale of the adjustment.

Ofwat's decision on the choice of which company to use as the benchmark for a notional efficient company involved an upper quartile company for draft determinations. But for final determinations Ofwat used more demanding benchmarks: the fourth-placed company in water and third-placed company in wastewater. Ofwat moved away from using the historical upper quartile stating concerns that this did "not appear to deliver a strong challenge for the sector",[1] in light of comparisons of modelled costs (subject to an upper quartile adjustment) relative to companies' updated business plan cost forecasts over the 2020-25 period. So, while there was no mechanistic link between wholesale base cost allowances and companies' business plan forecasts, it seems that these costs forecasts influenced the process to set allowances.

The CMA's approach to the notional company catch-up adjustment at PR19

In the PR19 references to the CMA, the focus was on wholesale rather than retail controls, so the CMA did not consider Ofwat's approach to residential retail cost assessment.

In relation to wholesale controls, the CMA adopted the following position on the notional company catch-up adjustment to be applied to results from wholesale base-plus econometric modelling:[2]

> *"Our cost models estimate how much it would cost the averagely efficient water company to cover base operations. However, we want to set cost allowances for a water company that is more than merely averagely efficient, and so we apply a 'catch up' efficiency challenge. Our decision is to use the company at the upper quartile as the benchmark and reduce the Disputing Companies' allowances accordingly. We consider this sets a challenging benchmark whilst acknowledging the limitations of our econometric modelling (and the consequent risk that the company will have insufficient allowed revenue to ensure a base level of service). Our benchmark is set at a similar, although slightly less demanding, level to Ofwat's."*

[1] Ofwat (2019) *PR19 final determinations: Securing cost efficiency technical appendix*, page 32.

[2] CMA (2021) *Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations: final report*; page 19.

In explaining its decision on this matter, the CMA explicitly listed a number of factors that it had placed "little or no weight on" but which had been raised as part of the appeals process. In particular, the CMA explained that it placed little or no weight on the following:[3]

> *… a comparison of the companies' business plans with the modelled allowances. We found it was more appropriate to set the efficiency challenge based on our assessment of the quality of the econometric modelling, rather than to seek specific outcomes."*

The CMA, therefore, seems to have explicitly decided against following Ofwat's approach of using comparisons with business plan forecasts to inform the notional efficient company catch-up adjustment, at least for wholesale costs which were the focus of its determination.

Use of cost forecasts in Ofwat's cost assessment for wholesale enhancements

The way that Ofwat used companies' forecasts of residential retail costs at PR19 was novel. But this is not the only area where Ofwat has used direct benchmarking of companies' cost forecasts in setting price control allowances. For enhancement expenditure, Ofwat's established practice has involved benchmarking of companies' cost forecasts.

At PR19, for a subset of enhancement categories – more than 50% of enhancement expenditure allowances – Ofwat's cost assessment involved cross-company benchmarking (including in some cases, unit cost benchmarking applied as deep dive assessments):

- For most of these categories, Ofwat's cross-company benchmarking used business plan forecasts of enhancement expenditure rather than historical enhancement expenditure.
- For the first time sewerage enhancement category, Ofwat's assessment involved triangulation of results from (i) econometric model estimated on forecast data; and (ii) econometric model estimated on historical data. For metering enhancement costs, Ofwat's benchmarking was focused on econometric models which drew on a data set covering both historical data and forecast data.

There are some points of difference to highlight for wholesale enhancement expenditure. In particular:

- In most cases, Ofwat did not have corresponding historical expenditure data that it could have used in place of company forecasts. Ofwat has not historically required companies to report outturn expenditure on enhancement activities with the same degree of granularity and even where there is granularity on costs there is no corresponding granularity on outputs or enhancement benefits.
- Ofwat's cost assessment practice, over many price controls, has placed much greater weight on companies' expenditure forecasts – and its review of those – in the case of wholesale enhancement expenditure than for ongoing wholesale base costs.

[3] CMA (2021) *Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations: final report*; pages 231-232.

- In some cases, there may be a view that enhancement activities involve doing quite different things from those done in the past, and that the historical evidence on past enhancements is of limited relevance relative to forecasts about future activities.

Position on the use of company cost forecasts in Ofwat's final methodology for PR24

In its PR24 final methodology, Ofwat said that it intends to use a combination of historical and (where appropriate) forecast efficiency evidence to set the catch-up efficiency challenge at PR24.[4] In the case of residential retail costs, this would imply a continuation of the type of approach used at PR19.[5]

There is also the possibility that for PR24 Ofwat uses companies' business plan forecasts as part of the input data used to estimate econometric benchmarking models (e.g. an approach which draws on both historical and forecast costs for the input data). This is not something that Ofwat did at PR19 but, as summarised below, is an approach that Ofgem has used for energy network price controls. Ofwat said that it intended to "cautiously consider" using business plan forecast data in its base cost models for PR24.[6] In response to some concerns raised by companies, said that it would "*only consider using business plan forecasts in the base cost models if they are robust and not significantly impacted by different company risk appetites*".[7]

Ofgem's use of companies' cost forecasts as inputs to econometric models

In recent price control determinations for gas distribution companies, and electricity distribution companies, Ofgem's cost assessment has involved econometric benchmarking models that draw on a sample of input data covering both outturn costs for previous years and companies' cost forecasts for future years.

In the case of the RIIO-GD2 price control for gas distribution companies, Ofgem's cost assessment drew in part on an econometric benchmarking model for totex with the following features:[8]

- The input data used to estimate the models included both historical expenditure data (over the period 2013/14 to 2019/20) and expenditure forecasts provided by companies in their price

---

4 Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24: Appendix 9 – Setting expenditure allowances*, page 36.

5 As explained earlier in section 2, at PR19 Ofwat applied an adjustment to modelled costs (derived from the application of econometric models to historical data) which was calculated by giving a 50% weight to an upper quartile efficiency adjustment based on companies' cost forecasts and a 50% weight to an upper quartile efficiency adjustment based on historical costs.

6 Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24 Appendix 9 – Setting expenditure allowances*, pages 14-15.

7 Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24 Appendix 9 – Setting expenditure allowances*, page 15.

8 Our understanding of Ofgem's approach and position for the GDN benchmarking is based partly on Ofgem's GD2 final determinations and partly on information reported in the CMA decision in the RIIO-2 appeals concerning SGN ground 4 on the efficiency benchmark: see CMA (2021) *Cadent Gas Limited, National Grid Electricity Transmission plc, National Grid Gas plc, Northern Gas Networks Limited, Scottish Hydro Electric Transmission plc, Southern Gas Networks plc and Scotland Gas Networks plc, SP Transmission plc, Wales & West Utilities Limited vs the Gas and Electricity Markets Authority Final determination Volume 3: Individual Grounds*

control business plans for the period up to 2025/26. This provided seven years of historical/outturn cost data and six years of forecast data.

- Ofgem's econometric models included two time trend explanatory variables: one over the full sample of historical and forecast data, and one which only applied in the period from 2020/21 onwards. This allowed for a different time trend for forecast costs compared to historical costs.
- Ofgem's modelling results indicated a negative time trend (cost reductions over time on a CPIH-adjusted basis) for the period of historical data and a positive time trend for the period based on companies' cost costs (cost increases over time on a CPIH-adjusted basis and higher costs in the forecast period compared to the historical period).
- Ofgem set a relatively demanding catch-up assumption (85th percentile, as opposed to the upper quartile at GD1) based on consideration of a range of factors including its view that the length of the sample period (e.g. longer than for the GD1 modelling) increased Ofgem's confidence in the results of the model (this was not the only consideration).

Ofgem took the position that using both historical expenditure data and expenditure forecasts provided by companies for the model input data increased the time series of data used for the model and, in turn, increased Ofgem's confidence in the results of the model. Ofgem said that further years of data should increase the accuracy of the modelled results and were apt to improve confidence.

Using companies' cost forecasts in the way it did also enabled Ofgem to adopt the position that its cost benchmarking took account of a number of changes over time affecting the gas distribution companies that might mean that their costs are higher in the future than the past. Some of these changes might be things that would be treated as enhancement costs in the context of water company price controls.

While some specific aspects of Ofgem's benchmarking approach were challenged during the RIIO-2 appeals to the CMA (e.g. the use of an 85th percentile efficiency benchmark), the principle of using cost forecasts as input data to the econometric models was not one of them.

3: Comparisons of cost forecasts at PR19 and outturn costs

In this section, we present comparisons of water companies' PR19 forecasts for residential retail costs against outturn costs for the first two years of the price control period. This section is organised as follows:

- Comparisons of total residential retail costs against companies' PR19 forecasts.
- Comparisons for subcategories of costs within residential retail costs.
- Companies' forecasts of CPIH inflation.

This section provides further context for the discussion and review in section 4 concerning the use of companies' cost forecasts in setting price control allowances for residential retail costs.

Some caution is needed because the analysis here is based on only two years of data rather than the full five-year period over which the business plan forecasts applied. Furthermore, some companies' forecasts may be intended to represent what they would expect to spend if they were operating at some efficient level (e.g. upper quartile on the benchmarking analysis) rather than what they will spend in practice. And companies' actual expenditure may be influenced by the final price control allowances set by Ofwat, which were not available at the time of the forecasts. Nonetheless, we consider the analysis presented in this section to be relevant in considering the appropriate approach to take at PR24.

It was not within the scope of this report to look at business plan forecasts at PR14, relative to outturn costs, but this might shed additional light on companies' past performance in forecasting residential retail costs.

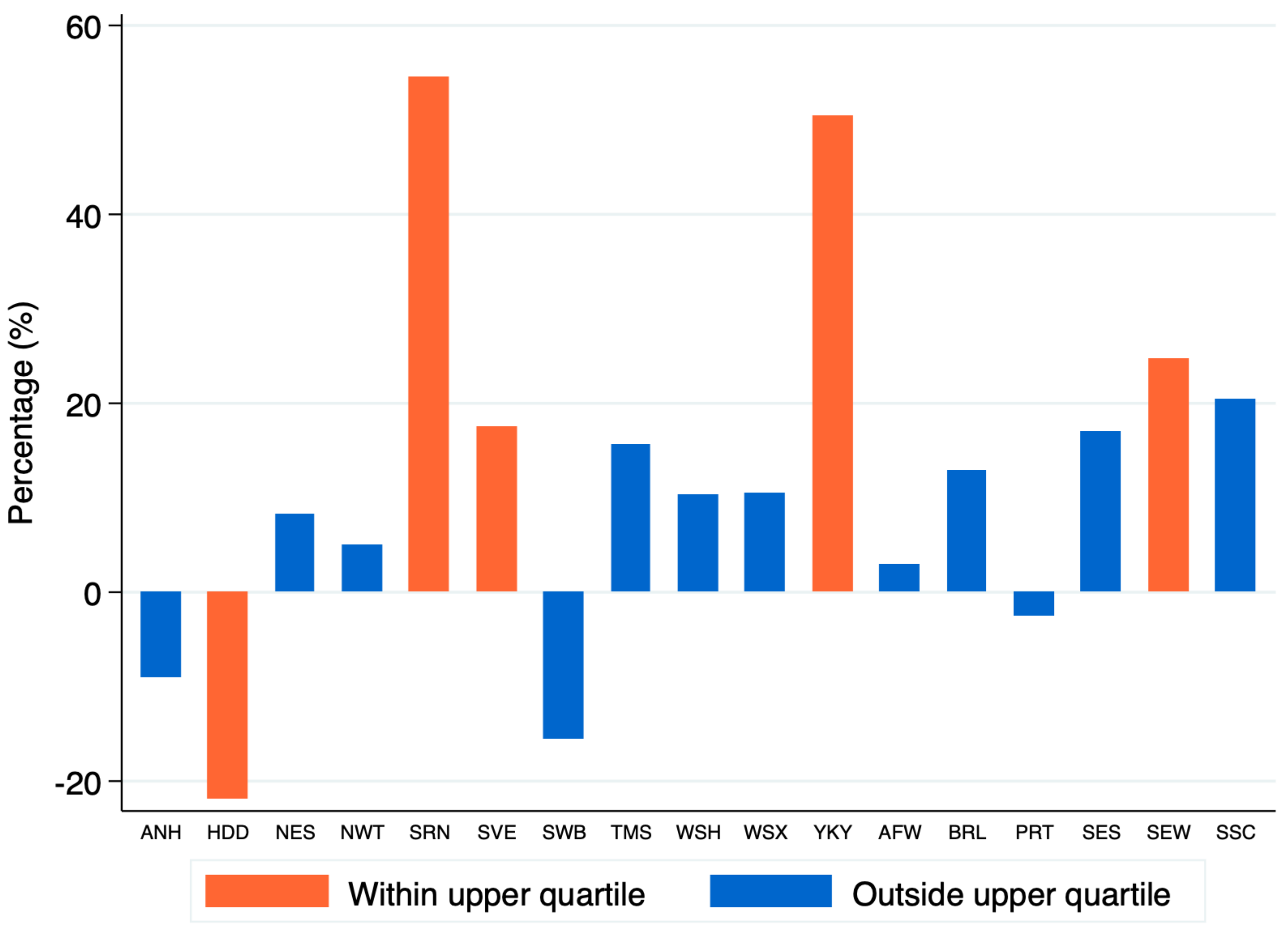
Comparisons of total residential retail costs against companies' PR19 forecasts

We have compared companies' PR19 business plan forecasts for residential retail costs against outturn costs, using the data available to date. Figure 1 shows how companies' outturn costs across 2020/21 and 2021/22 in aggregate compare with the costs as forecast in companies' business plans. For this analysis, we have drawn on the data set out in the Excel files published by Ofwat as part of its final determinations. These reflect any revisions made by companies over the course of the PR19 process and so the forecasts do not necessarily match those from companies' original business plans. We show in red bars those companies that Ofwat considered to be within the upper quartile in terms of the efficiency of their cost forecasts.

As shown in the figure, taking 2020/21 and 2021/22 together, companies' outturn costs across tended to be above those that they had set out in their PR19 business plans. On average, across companies, outturn costs were 12% above business plan forecasts.

Across the set of companies, Hafren Dyfrdwy, South West Water and, to a less extent, Affinity Water and Portsmouth Water, stand out as those whose outturn costs were below the levels that the companies had forecast. At the other end of the spectrum, outturn costs for Yorkshire Water and Southern Water were far greater than the level set out in those companies' business plans, more than 50% above forecasts.

Figure 1 Total retail costs over- or under-spend compared to PR19 allowances (2020/21 and 2021/22)



Comparisons for subcategories of costs within residential retail costs

Companies' commentaries in their APRs for these years raise the Covid-19 pandemic as a factor driving some of the higher costs in those years, particularly in 2020/21.

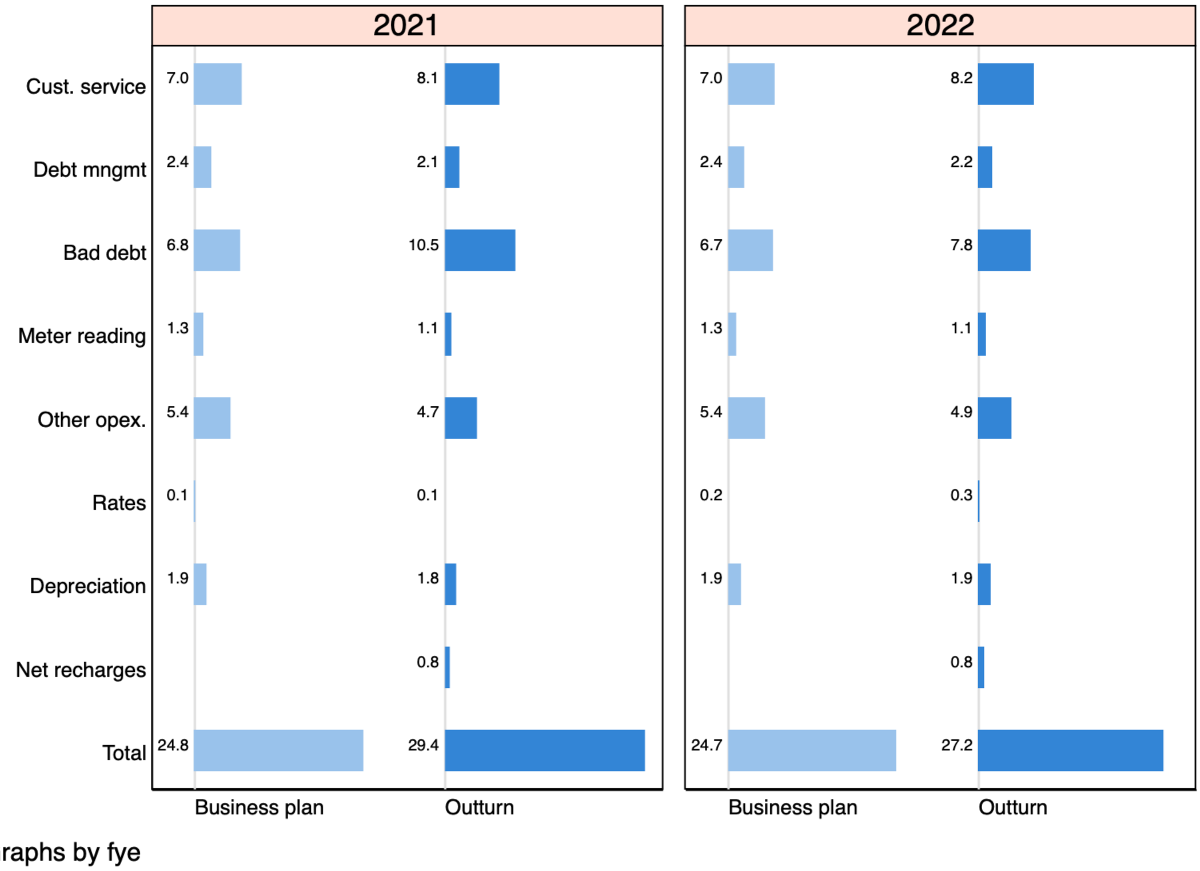
We have compared outturn costs with companies' forecast costs at a more granular level, at the level of individual cost lines in companies' APR reporting. Figure 2 shows that comparison. In the chart, costs are expressed in nominal terms, on a per connected household basis, and values have been averaged across companies.[9]

Figure 2 suggests that, at an industry-wide level, the cost lines that have diverged the most, in absolute terms, from the values forecast in business plans were costs related to customer service and bad debt costs. In 2020/21, for example, companies' customer service costs were, on average £1.10 per household greater than what they had forecast (£8.10 compared to £7.00 per household) and bad debt costs were £3.70 per household higher (£10.50 compared to £6.80 per household).

[9] For this chart, we drew on data included in the Excel files companies submitted in January 2019 as part of their revised plans (from their initial September 2018 submissions). We are not aware that data from later versions of the plans containing such granular data were published. Data for Thames Water are not included in the analysis underlying the charts as such granular data were redacted from that company's submission.

On the other hand, meter reading costs were £0.20 and 'other operating costs" around £0.6 per household below business plan forecasts, in each of the two years.

Figure 2 Out-turn and business plan costs (2020/21 and 2021/22)

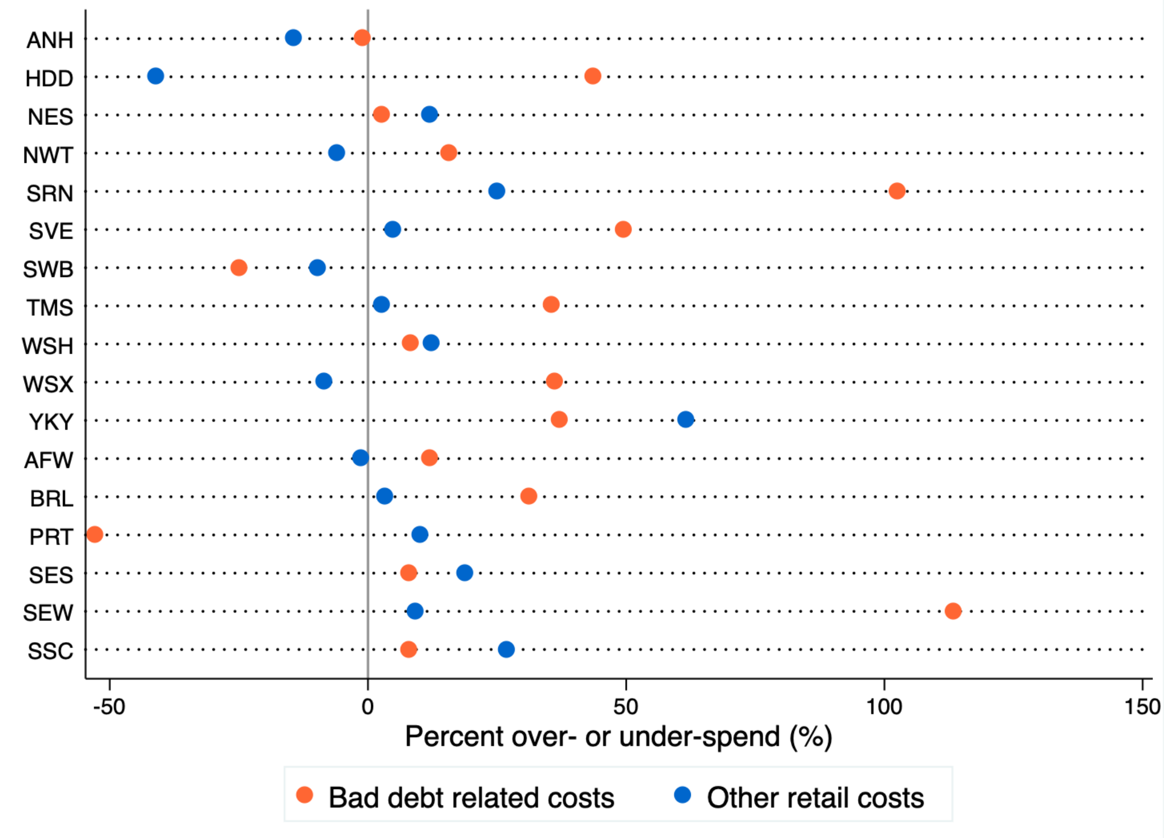


Because Figure 2 is at an industry-average level it hides the variation that is observed across companies. Figure 3 below shows some of the variation across companies. For presentational reasons we have grouped together the cost lines falling within each of the two categories of costs for which Ofwat developed separate models at PR19: bad debt related costs, and other retail costs.[10]

The figure echoes the general pattern set out in the comparisons in previous charts, namely that outturn costs tended to be above those forecast in companies' PR19 business plans, though there is a difference between the two cost categories. In respect of bad debt related costs, companies over-spent compared to their business plan forecasts by, on average, 25%, whilst the average over-spent in relation to other retail costs by around 6%. This is consistent with the view that the Covid-19 pandemic contributed substantially to the discrepancy between forecast and outturn costs.

[10] To construct this figure, we have been able to draw on business plan figures included within the set of files published by Ofwat at final determinations, as these included totals for the two broad cost categories. For presentational reasons, the figure shows the average across 2020/21 and 2021/22 of the difference between outturn and forecast costs, rather than showing figures for the two years separately.

Figure 3 Over- or under-spend compared to business plan forecast costs (2020/21 and 2021/22)



Companies' forecasts of CPIH inflation

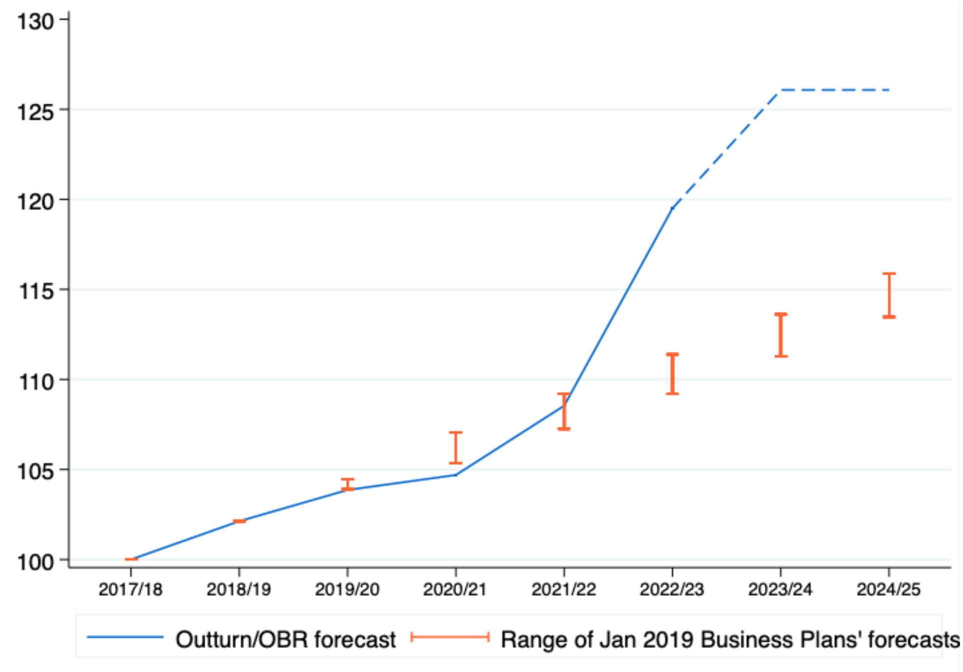
At PR19, and in respect of residential retail costs, companies provided figures on their forecast of costs over the AMP7 years in nominal terms. As part of the information included in the business plan Excel templates, companies also provided information on their forecast of RPI and CPIH inflation over AMP7.

It is of interest to contrast companies' forecast of inflation for that period with outturn inflation rates for 2020/21 and 2021/22, and with more recent forecasts of inflation for the later years. Differences between companies' forecasts at the time of developing their business plans and outturn inflation could help explain some of the observed differences between costs forecast in business plans and outturn costs.

Figure 4 shows outturn CPIH index in the period from 2017/18 to 2021/22 and an OBR forecast for CPI for the years to 2024/25.[11] The figure also shows, for each year, the range across companies' forecast of CPIH inflation reported in their January 2019 Business Plan submissions.

[11] The OBR forecast is based on its November 2022 forecast for CPI.

Figure 4 Outturn and forecasts of CPIH index (2017/18=100)



The figure shows that companies' forecasts of the CPIH index for 2020/21 were slightly above outturn CPIH and that companies' forecasts for 2021/22 were roughly in line with outturn. This chart suggests, therefore, that substantial gaps between outturn costs and companies' forecast costs for 2020/21 and 2021/22 – as shown earlier in this section – are not driven primarily by economy-wide inflation. Put differently, the discrepancies show above between companies' forecasts and their outturn costs in 2020/21 and 2021/22 do not seem to be the result of companies making inaccurate forecasts of CPIH inflation in the period to 2021/22.

From 2022/23 to the end of AMP7, companies' forecasts of CPIH at PR19 are well below the level that has been forecast for that period by the OBR. If the CPIH does evolve as forecast by OBR, this would bring the potential for even greater variation between outturn costs and companies' PR19 forecasts than we have seen above.

4: Review of the case for using forecast cost data at PR24

In this section we consider and discuss a number of issues which relate to the case for Ofwat using water companies' forecasts of residential retail costs in settling retail price control allowances at PR24. We take the following topics in turn:

- Our view on the rationale for Ofwat's use of forecast at PR19.
- Overview of concerns about Ofwat's use of cost forecasts at PR19.
- Limitations in Ofwat's business plan assessment for retail cost efficiency at PR19.
- Interactions with the use of upper quartile benchmarks.
- The relevance of internal assurance processes to the credibility of cost forecasts.
- Insights from comparisons with competitive markets.
- Companies' accountability for their cost forecasts.
- Mitigating risks associated with the use of company forecasts at PR24.
- Use of cost forecasts as input data to the econometric benchmarking models.

Our main interest is in the potential use of water companies' forecasts of residential retail costs in a similar way as at PR19: to set an adjustment that is applied to modelled costs across the industry as a means to project cost benchmarks over the forthcoming price control period. This seems the most likely way that Ofwat would use forecasts at PR24. In the final subsection above we consider the potential use of forecast costs as input data to its econometric benchmarking models.

Our view on the rationale for Ofwat's use of company forecasts at PR19

We can understand why at PR19 Ofwat was keen to try to take account of business plan cost forecasts via its adjustment to modelled costs, and why it has given itself a clear option to do so similarly at PR24.

There is a potential role for using companies' business plan forecasts as a check on the retail cost allowances that Ofwat is considering setting, and there is also a potential role for these as a more direct input when projecting cost benchmarks into a forthcoming price control period.

On the first of these potential roles, the following points are particularly relevant:

- As a general point, we would expect a regulator such as Ofwat to be concerned about setting allowances based primarily on benchmarks derived from historical costs, in a context where such allowances would, for a number of companies, be considerably more than the cost allowances that companies are seeking through their business plans.
- This type of concern seems to be particularly relevant given indications of quite a significant pace of improvement in residential retail cost efficiency over time. Ofwat said that: "*The fact that*

the majority of companies submitted forecasts that are significantly more efficient than historical expenditure is evidence of the pace at which this service is transforming".[12]

- The risk of setting allowances at PR19 that were too high was exacerbated by some features of Ofwat's approach to residential retail cost assessment. First, its econometric benchmarking models were largely static in the sense that they did not allow for any changes in costs over time (other than if cost driver explanatory variables changed). Ofwat's models essentially imposed an assumption that, leaving cost driver effects aside, unit costs were constant *relative to CPIH* over the historical sample period – despite this assumption being at odds with the evidence available for that period. Second, the retail cost assessment that Ofwat presented in its PR19 final determinations did not include any analysis (outside of the use of company forecasts) of historical trends in water company unit costs or analysis of the effects of ongoing productivity and input price changes on efficient retail costs, which reinforced the static nature of the analysis.
- We doubt that Ofwat would have placed the same weight on business plan forecasts had these led to higher retail cost allowances than the figures based on a more conventional approach of a notional efficient company adjustment calculated using historical data combined with adjustments for ongoing productivity improvements and input prices. Ofwat's PR19 methodology was quite vague about how it would take account of business plan forecasts. As with many previous price control reviews, we would expect Ofwat to generally find it straightforward to set allowances that are below the forecasts submitted by companies, at least in the absence of good evidence from companies to substantiate their higher forecasts.

The points above concern the possible role of business plan cost forecasts as a means for Ofwat to help check that its proposed cost allowances are not too high.

Furthermore, as we have seen from Ofwat's approach to residential retail costs at PR19, there is the possibility of using business plan cost forecasts more directly as an input to the calculations that are used to set allowances.

In the absence of using business plan cost forecasts, a more conventional approach to projecting cost benchmarks into a forthcoming price control period would be to make explicit assumptions about the opportunities for ongoing productivity improvements (or frontier shift efficiency) and changes over time in input prices (or assumptions on efficient rates of changes in unit costs, representing the combined effects of productivity and input price effects) and use these assumptions to adjust the cost benchmarks that were derived from the historical analysis. This type of approach tends to be based heavily on historical evidence, rather than forward-looking evidence, and can often rely on evidence from completely different parts of the economy (e.g. productivity data on the UK as a whole or for specific broad UK sectors within it).

One of the main benefits that we see in using business plan forecasts, in the way that Ofwat did at PR19, is that it provides an alternative method for projecting cost benchmarks over a forthcoming price control period in a way that uses evidence that is focused on water companies' residential

12 Ofwat (2019) *PR19 final determinations: Securing cost efficiency technical appendix*, page 128.

retail activities and which can be more forward-looking.[13] How forward-looking companies cost forecasts are in practice depends on how companies have made them (e.g. the forward-looking perspective could be an illusion if companies' forecasts are based on simple extrapolations of historical evidence).

As discussed later in this section, we consider that there is a stronger case for using this type of evidence to project cost benchmarks alongside other types of evidence, rather than in place of these. This reflects the reality that all of the approaches have significant downsides and there are benefits in this context in drawing on multiple approaches.

Overview of concerns about Ofwat's use of cost forecasts at PR19

Despite the potential rationale set out above, we see significant concerns with using business plan forecasts in the way that Ofwat did at PR19. These concerns relate in particular to the risks that the cost forecasts that companies include in their business plan may involve considerable levels of inaccuracy.

As indicated in section 3, the evidence from the 2020-25 period so far suggests that, even allowing for the effects of Covid-19, some companies' cost forecasts were well below what they have spent in practice, which highlights the risk that business plan cost forecasts may be too low (alongside the more familiar regulatory risk that companies' cost forecasts are too high).

There are a number of factors which may act to limit the accuracy of companies' business plan cost forecasts (some of these are interrelated):

- **Uncertainty about future costs**. There is considerable uncertainty about future levels of efficient residential retail costs over a forecast horizon spanning five or more years.
- **Prioritisation of forecasting effort**. Retail is a relatively small part of the water company revenue being determined at the price review, and companies may not give sufficient thought and attention to producing high-quality forecasts of future retail costs amidst the demands of the wide-ranging price review process (especially in a context where companies expect their retail price controls to be based on Ofwat's cross-company benchmarking rather than each company's own forecasts).
- **Accountability for accuracy of cost forecasts**. The accuracy of cost forecasts may be adversely affected by companies lacking clear accountability for their forecasts – or a lack of obvious skin in the game. This may be the case where Ofwat's approach is based primarily on benchmarking analysis (and hence driven primarily by the costs incurred and perhaps forecast by other companies) and whether a company submitted an accurate or inaccurate forecast of its own retail costs would generally be expected to have limited effects on the level of retail price

[13] Another alternative to what Ofwat did at PR19 would be to adopt an approach of capping each company's allowance as the maximum of the allowance derived from Ofwat's benchmarking analysis and the company's own cost forecast. But this type of approach risks providing (or exacerbating) financial incentives for companies to submit inflated cost forecasts as part of their business plans. For retail at PR19, Ofwat used business plan forecasts in its cost assessment in a way which was less likely than this to adversely affect incentives.

controls its faces over the period. We discuss this issue later in the section; it is perhaps less of a concern at PR24 than at PR19, but still a relevant issue.

- **Incentives that may influence or distort cost forecasts (1): over-estimation of costs**. Typically, UK regulators such as Ofwat and Ofgem have adopted approaches to cost assessment that are designed to help guard against the risks that companies' cost forecasts for the price control period are too high. Companies may have financial incentives, especially in a context of uncertainty, to err on the side of submitting forecasts which might be too high rather than too low, and may face limited incentives to provide forecasts that reflect strong performance on cost efficiency.
- **Incentives that may influence or distort cost forecasts (2): under-estimation of costs**. It is also possible that companies face incentives to submit business plan forecasts that are lower than their central forecast and which imply levels of efficiency improvement which are unlikely to be achievable. This scenario might arise, in particular, from a combination of the financial, procedural and reputational incentives created by Ofwat's business plan assessment process which rewards companies that the regulator views as having submitted ambitious or relatively efficient cost forecasts and/or penalises companies who the regulator views as having submitted unambitious or relatively inefficient cost forecasts. At PR19, Ofwat's business plan assessment process was designed to give better grades to companies that forecast lower costs, rather than seeking to reward the accuracy or quality of forecasts. Ofwat's PR19 business plan assessment process may have provided at least some companies with an artificial incentive to submit business plan forecasts for retail costs that were lower than their own central view of their efficient costs over the 2020-25 period. For PR24, Ofwat plans a similar form of business plan assessment process with of use of financial, procedural, and reputational incentives. The business plan incentives are likely to more influential the less accountable a company is for the actual accuracy of its forecasts.

At this stage, and in the context of PR24, our view is that:

- The net impact of the factors above on companies' business plan cost forecasts is difficult to gauge, and likely to vary across companies' depending on their business strategy and their subjective views about the influence that different elements of their business plan will have on both the business plan assessment process and the setting of their allowances.
- There is no reason to be confident that the net impact of incentives and accountability will mean that PR24 business plan forecasts will be *accurate* or even *reasonably accurate.*

The way that Ofwat used business plan forecasts at PR19 for retail cost assessment meant that any inaccuracy in one specific company's forecasts did not just affect that company or its customers, but could affect all other companies and their customers. Relatively reliable forecasts could be crowded out by unreliable forecasts. We discuss this further in the section further below on "Interactions with the use of upper quartile benchmarks".

Interactions with the CMA position on the wholesale catch-up adjustment

As summarised in section 2, in its determinations of wholesale controls in the PR19 references, the CMA decided against following Ofwat's approach of using companies' business plan cost forecasts – and comparisons of these against modelled costs – to choose the efficiency benchmark for the notional efficient company catch-up adjustment (e.g. upper quartile or a more demanding benchmark).

The CMA seems to have shown greater caution than Ofwat in the use of companies' cost forecasts in setting price control allowances. This caution may reflect some of the concerns above about the accuracy of companies' business plan cost forecasts.

However, the way that Ofwat used companies' cost forecasts for residential retail cost assessment differed in some significant ways from its approach at the wholesale level. In particular, Ofwat's approach to residential retail can be seen to take account of (some) companies' views on changes over time in the efficient levels of retail costs which arise from ongoing productivity improvements and input price effects, and it goes beyond a more static upper-quartile adjustment as used by the CMA for wholesale base costs. For the type of approach used for residential retail, we doubt that the CMA would reject the use of company forecasts as a matter of principle. We would instead expect the CMA to pay more attention to what steps Ofwat has taken to avoid using unreasonable or misleading forecasts and whether it has placed full reliance on company forecasts or triangulated across different types of evidence and approaches.

Limitations in Ofwat's business plan assessment for retail cost efficiency at PR19

In its structured assessment of water companies' business plans at PR19, Ofwat had a question focused on residential retail cost efficiency, which was: "*How well evidenced, efficient and challenging are the company's forecasts of retail expenditure, including bad debt costs?"*

Although this question referred to how well evidenced a company's forecasts were, Ofwat's approach to grading companies' business plans meant that there was quite a mechanistic relationship between Ofwat's assessment and its perceptions of the efficiency of companies' forecasts. Ofwat explained its approach as follows:[14]

> *"Our grade for each question is calculated as follows: A: company totex is more than 5% more efficient than our view of totex; B: company totex is up to 5% less efficient [than] our view C: company totex is up to 20% less efficient than our view; and D: company totex is more than 20% less efficient than our view."*

This means that, on Ofwat's own admission, its approach to grading companies' business plans in relation to residential retail costs focused on the level of those forecasts - and how these compared to results from Ofwat's cost benchmarking - without regard to the evidence provided to support the forecasts or the credibility of those forecasts.

Furthermore, Ofwat highlighted the following as examples of good practice from its assessment of companies' forecasts for residential retail costs:[15]

> *"Southern Water provide a high quality plan and ambitious cost forecasts for its residential retail service. The plan sets out its 'retail transformation programme' for PR19. For Southern Water this plan and cost forecasts show a commitment to change.*

---

14 Ofwat (2019) PR19 initial assessment of business plans: Summary of test area assessment, page

15 Ofwat (2022) *PR19 initial assessment of business plans: Summary of test area assessment*, page 89.

Yorkshire Water submits the most efficient plan in residential retail, continuing a track record of efficiency in this service."

Ofwat seems to have equated a high quality plan with ambitious cost forecasts. As seen in section 3, Southern Water has overspent its business plan forecasts by over 50% which raises serious questions over the quality of its plan. While part of the overspend may reflect the effects of Covid-19 on retail costs, which would not have been anticipated, this scale of the overspend is far higher for Southern Water than for most other companies.

Within the group of three companies who achieved the best grade for retail cost efficiency in Ofwat's PR19 IAP assessment, the average over-spend so far in the 2020-25 period has been far higher than for other companies, as shown in the table below.[16] Each of these companies had a higher-than-average over-spend and two of them (Southern Water and Yorkshire Water) had extremely high levels of overspend (over 50%). This provides some evidence that, where companies submitting residential retail cost forecasts that Ofwat found to be "ambitious" at PR19, these were also companies who had underestimated their residential retail costs in the 2020-25 period.

Table 1 Comparison of outturn over-spend vs forecasts against IAP grade for retail cost efficiency

Ofwat assessment grade for test area CE3 (retail cost efficiency)	Number of companies receiving the grade	Average over-spend vs business plan forecast (2020-2022) for grade
A	3	41%
B	5	6%
C	5	6%
D	4	12%

Some caution is needed because the analysis here is based on only two years' data rather than the full five-year period over which the business plan forecasts applied. Furthermore, some companies' forecasts may be intended to represent what they would expect to spend if they were operating at some efficient level (e.g. upper quartile on the benchmarking analysis) rather than what they will spend in practice. Nonetheless, the scale of differences across companies of different grades is quite striking and does cast doubt about the reliability of forecasts that seem to be the most ambitious and for which Ofwat awarded the best grades at PR19.

Interactions with the use of upper quartile benchmarks

The risk of Ofwat using forecasts that are too low is exacerbated insofar as Ofwat's approach involves putting greater weight on the forecasts that it calculates as being at an upper-quartile, or similar, in terms of the relative efficiency of forecasts across the industry (rather than using a

[16] The comparison of forecast and outturn costs is based on the company forecasts as set out in the Excel files published by Ofwat as part of its final determinations. These reflect any revisions made by companies over the course of the PR19 process and so the forecasts do not necessarily match those from companies' original business plans as assessed at IAP stage.

forecast ranked at the median in terms of efficiency or making no adjustment for catch-up efficiency improvements and thereby setting allowances based on industry-average efficiency levels).

We expect that there will *be variation across companies* in terms of both:

- The direction of any inadvertent forecasting error (e.g. whether inadvertent forecasting error leads to over-estimation or under-estimation of costs).
- Whether forecasts are materially influenced by (or distorted by) incentives arising from the price control framework and price review process and, if so, whether this acts to increase or decrease the forecasts submitted.

In the context of this variation, the cost forecasts that Ofwat identities as being upper quartile (or similar) in terms of their perceived cost efficiency are, on the balance of probability, likely to reflect a greater degree of under-estimation of costs than applies on average across the forecasts submitted by all companies. This is not to say that the upper quartile cost forecast will tend to be an under-estimate of efficient costs over the forthcoming price control period, just that it is more likely to be an under-estimate than the average forecast.

To take an example of what this means in practice, if five companies out of 17 (i.e. less than 30%) under-estimate future costs and all the other companies over-estimate future costs, then an approach of Ofwat using the forecasts of the upper quartile company to adjust modelled allowances for all companies would, on the balance of probability, tend to involve overly-demanding efficiency adjustments.

Based on two years of outturn data for the 2020-25 period so far, we found that the average over-spend for the companies ranked by Ofwat as upper quartile (or better) in terms of their business plan forecasts was 25%, compared to 9% for the other companies.

This issue is relevant when considering risks around the use of companies' business plan forecasts to set allowances and the choice of efficiency benchmark.

There may be a way for Ofwat to mitigate, to some degree, this specific concern about the greater forecasting inaccuracy from an upper quartile forecast. This is as follows:

1. Ofwat could first calculate a more conventional upper quartile efficiency adjustment based on historical costs.
2. Ofwat could then calculate a further adjustment for dynamic effects (e.g. ongoing productivity improvements and changes to nominal input prices) calculated as the *average* or *median* across companies of each company's ratio of (a) forecast costs in nominal terms to (b) modelled costs (without applying any efficiency adjustment to these modelled costs).
3. To make projections for the 2025-30 period, both the historical upper quartile adjustment and the further adjustment for dynamic effects would be applied to modelled costs.

This approach would enable Ofwat to retain the use of an upper quartile efficiency benchmark (if appropriate), while potentially improving the way that company forecasts are used to bring a forward-looking perspective into the projection of cost benchmarks.

Companies' accountability for their cost forecasts

We would be particularly concerned about a company's retail cost forecast being used in setting allowances for all companies if the company limited accountability for that forecast or lacks a sufficient stake in its accuracy.

To take an extreme example, suppose a company thought that there was no likelihood that its business plan cost forecast would affect its own cost allowances, and instead treated the forecast as simply a means to improve its performance in Ofwat's business plan assessment process (e.g. by demonstrating greater ambition or cost efficiency), or to gain reputational benefits with Ofwat or other stakeholders. In these hypothetical circumstances, the forecast would not seem to provide relevant evidence on future levels of efficient retail costs and it would not be reasonable to use this forecast to set more challenging retail cost allowances for other companies.

There may have been incidences of this type of scenario arising in some areas at PR19, given the more limited role that companies' forecasts had played at PR14.

However, for PR24 we do not consider that this extreme scenario is likely to apply. At PR19 Ofwat used companies' forecasts directly in setting the retail cost allowances, explicitly calculating an adjustment that was applied to modelled costs for all companies based on comparisons of companies' business plan forecasts against modelled costs. In its PR24 final methodology, Ofwat has not indicated any move to play less weight on evidence from companies' business plan forecasts than at PR19. Given this, companies should be well aware that their own forecasts could influence their allowances, and that submitting an unduly low forecast could lead to unduly low allowances.

It seems probable that, in light of the approach used by Ofwat at PR19, and the experience of over-spends relative to forecasts across the industry so far, water companies take steps to improve the accuracy of their forecasts of residential retail costs at PR24.

Nonetheless, even though companies have some stake in their forecasts, this is not enough to ensure that these provide sufficiently reliable evidence for price control cost assessment.

The relevance of internal assurance processes to the credibility of cost forecasts

There is a potential argument that the types of assurance processes and requirements for board-level approval needed for price control business plans mean that these forecasts can be treated as reliable and credible (at least insofar as any forecast of the future can be). Indeed, Ofgem seems to have adopted this position during the appeals to the CMA on the RIIO-2 price controls. The CMA reported as follows:[17]

[17] CMA (2021) *Cadent Gas Limited, National Grid Electricity Transmission plc, National Grid Gas plc, Northern Gas Networks Limited, Scottish Hydro Electric Transmission plc, Southern Gas Networks plc and Scotland Gas Networks plc, SP Transmission plc, Wales & West Utilities Limited vs the Gas and Electricity Markets Authority Final determination - Volume 3: Individual Grounds*, paragraph 9.76.

"GEMA said that it had confidence in the forecast costs provided because all network companies were required to provide assurance reports and board-level sign-off on the final RIIO-GD2 business plans and BPDTs."

This comment was made in relation to a specific point of appeal, and may not reflect the general position of Ofgem in broader areas of cost assessment. Nonetheless, it is a useful example of the type of argument that might be made in relation to the use of company forecasts and we comment below on this potential argument in the context of Ofwat's cost assessment.

Ofwat's price review process would be far quicker and cheaper if assurance processes and board sign-off meant that it could have confidence in business plan cost forecasts and just plug these into the calculation of price control allowances. However, this is clearly not how it works in practice, with Ofwat expending substantial resources on cost assessment processes that review and often displace company forecasts. This is evidence that Ofwat lacks confidence in the business plan forecasts, despite the requirements for company assurance processes and board-level approval.

These requirements can help to improve the reliability of the cost forecasts in companies' business plans, compared to a price review process without those requirements, but there is no reason to think that they are sufficient. This point applies just as much to the concern that a particular company's cost forecasts may be too low as it does to the more familiar concern that a company's cost forecasts may be too high.

On this basis, we do not consider that requirements for assurance processes and board-level approval remove the need for Ofwat to check the credibility of cost forecasts if they are to be used in a similar way to those for residential retail costs at PR19.

Insights from comparisons with competitive markets

While it should not determine how the details of economic regulation are applied, it can also be helpful to draw on insight into what happens in competitive markets when considering the role of cost forecasts in setting water company price controls.

We consider that the following features of competitive markets are particularly relevant to highlight:

- Companies set prices in light of their expectations of future costs, not simply on the basis of the costs they have incurred historically. Companies compete, in part, in terms of their forecasts of costs. Those companies with a more optimistic view about their ability to reduce or restrain their costs will tend to reflect this in their pricing decisions and, in turn, gain a competitive advantage over their rivals.

- Companies are able to adjust their prices as new information on costs is revealed over time and as their forecasts are updated over time. While companies' cost forecasts will feed into their pricing decisions and commercial success, it is not typically the case that companies are constrained by forecasts of costs over a five-year time horizon with no scope to adjust prices if those forecasts no longer fit with the latest expectations. Competitive markets normally allow for more frequent feedback processes, and the updating of prices in light of new information, than water company price controls.

- Companies that make bad forecasts of their future costs (e.g. worse forecasts than their competitors) will tend to suffer from this – with risks to their financial viability and ability to compete. This acts as a discipline and incentive on the quality of cost forecasts that companies make and as a market selection mechanism, such that companies whose forecasts of costs are consistently worse than their competitors will tend to lose market share and their influence in the market.

In the table below, we draw some implications of the features of competitive markets.

Table 2 Insight from competitive markets for Ofwat's use of forecasts in cost assessment

Feature of competitive market	Potential implications for retail cost assessment at PR24
Companies set prices in light of their expectations of future costs, not simply on the basis of the costs they have incurred historically.	Comparison with competitive markets suggests that – as a matter of principle – it is not unreasonable for price control cost allowances to be set in a way that reflects companies' forecasts of future costs, especially where these look competitive relative to analysis of companies' historical costs.
Companies are able adjust their prices as new information on costs is revealed over time and as their forecasts are updated over time.	In deciding what weight to give company forecasts in retail cost assessment, we should recognise that the time period of water companies' business plan forecasts brings a greater degree of forecasting error than the shorter-term and more adaptable cost forecasts that typically underpin prices in competitive markets.
Companies that make bad forecasts of their future costs will tend to suffer from this.	In deciding what weight to give company forecasts in retail cost assessment, we should recognise that there is not the same discipline on regulated companies in terms of the quality of their forecasts and that there are no market selection processes leading to the exit of companies that consistently make bad forecasts.

There are also potential insights for water companies' cost of capital from comparisons with competitive markets, in a context where weight is placed on forecast costs, but we take these as outside the scope of this project.

The second point in the table above is also relevant to the question of whether the residential retail price control should involve some form of inflation indexation or some other form of adjustment mechanism, to help protect companies and customers from uncertainty about economy-wide inflation and its impacts on efficient retail costs, in the context of five-year price controls.

Mitigating risks associated with the use of company forecasts at PR24

Ideally the price control framework, and Ofwat's approach to the price review process, would provide water companies with a combination of financial and reputational incentives that encouraged them to (a) engage in a proportionate amount of analysis and effort to develop good forecasts of future (efficient) costs, and (b) submit cost forecasts in their business plan which represent their central forecasts of what their costs would be (assuming they operated efficiently).

This is unlikely to be achievable in practice. But over time, regulators can adapt their approaches to help move things in this direction and there may be amendments to aspects of Ofwat's approach which could bring further improvements for PR24. These issues are outside the scope of this report.

Further to this, and turning to the specifics of residential retail cost assessment, we see two main opportunities to mitigate risks around the accuracy of company forecasts at PR24, beyond established practice:

- **Assessment of the credibility of company forecasts**. While Ofwat's established practice in cost assessment is geared towards addressing the risk that a company's cost forecasts are too high, its practice to date seems to have placed limited emphasis on the risks that a company's cost forecasts are too low. This risk becomes more important if one company's forecast may be used to calculate allowances for other companies, so that it is not just the company that forecasts too low which suffers from this.
- **Triangulation with approaches that do not rely on company forecasts**. Further mitigation can be achieved by Ofwat's cost assessment not giving 100% weight to methods reliant on companies' cost forecasts when setting allowances. For example, by setting allowances based on the average of cost projections based on an approach involving company forecasts and one or more approaches that are not reliant on these.

As we have seen above, at PR19 Ofwat's assessment of companies' residential retail cost forecasts seemed to focus on their apparent ambition and efficiency (under Ofwat's benchmarking analysis) rather than their credibility or the quality of evidence supporting them. In responding to Ofwat's PR24 final methodology, some stakeholders raised concerns of this nature. Ofwat summarised as follows:[18]

> *"some respondents considered that our quality assessment will place insufficient emphasis on the deliverability of proposals. As a result, they considered that a company could be encouraged to include in its plan, and ultimately be rewarded for, ambitious proposals that it is unlikely to deliver."*

In light of these concerns, Ofwat said that it would adapt its approach to business plan assessment:[19]

> *"For final methodology we include a new minimum expectation in our quality assessment for each company to provide sufficient and convincing evidence that it can credibly deliver the proposals in its plan. The amount of evidence we require from each company will be proportionate to the extent of the gap between its track record of performance, including delivering improvements, and the proposals in its plan. This minimum expectation will provide Ofwat and other stakeholders with greater confidence that plans will deliver for customers, communities and the environment and so supports the credibility of the sector overall."*

---

[18] Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24 Appendix 12 – Quality and ambition assessment*, page 5.

[19] Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24 Appendix 12 – Quality and ambition assessment*, page 8.

As far as we can tell, Ofwat has not set out the full implications of it finding that a company has not provided "sufficient and convincing evidence that it can credibly deliver the proposals in its plan". Nonetheless, as a matter of logic and good practice, we do not see how Ofwat could reasonably use the cost forecasts for such a company to determine or adjust the allowances set for other companies.

For instance, if it adopts a similar type of approach to that used at PR19, Ofwat might need to make calculations of upper quartile efficiency of companies' cost forecasts by only comparing the forecasts of companies that have met its minimum expectation on the credibility of business plan forecasts.

On this basis, Ofwat's inclusion of the new minimum expectation on the credibility of companies' cost forecasts should help to mitigate – to some degree at least – concerns that all companies' allowances are adversely affected by forecasts provided by some companies that lack credibility or do not have a good evidential basis.

A concern that companies might have is that in practice, it can be difficult and time-consuming for Ofwat to assess the credibility of companies' cost forecasts, to tackle risks that forecasts are too low, and that Ofwat may not put much resource into this aspect of its price review process. There may be a role for other water companies raising concerns, during the PR24 process, if a specific company does not seem to have met Ofwat's expectations for "sufficient and convincing evidence that it can credibly deliver the proposals in its plan" – especially where that companies' proposals on costs could affect the allowances Ofwat sets across the industry – and where this does not seem to have been detected by Ofwat.

A further perspective on the credibility of company forecasts might be obtained by looking in more detail at the different financial incentives that might influence company forecasts. For instance, this could involve some quantitative analysis to consider the scale of financial rewards from the business plan assessment process, and the degree of impact of retail cost forecasts on these, relative to the potential for a company's forecast to influence its own allowances via the cost assessment process. Such analysis is unlikely to be conclusive, but could help provide some tentative guidance on whether the forecasts provided by water companies as part of their business plans can reasonably be interpreted as forecasts of their (efficient) costs – rather than simply as numbers created in response to incentives created by the price review process.

Use of cost forecasts as input data to the econometric benchmarking models

Our main interest in this report is the potential use of water companies' forecasts of residential retail costs in a similar way as at PR19. This seems the most likely way that Ofwat would use forecasts at PR24.

We have also considered an approach in which companies' cost forecasts are used directly as part of the input data used to estimate econometric models. For example, as highlighted in section 2:

- Ofwat estimated econometric models using forecast cost data in its benchmarking of a number of wholesale enhancement expenditure at PR19. In most cases Ofwat's models used only forecast data but some of its modelling did use a combined data set and historical and forecast data and in other cases Ofwat triangulated with results from models estimated on forecast data.

- Ofgem estimated econometric models using a sample covering historical cost data and companies' cost forecasts in its benchmarking analysis for energy network price controls (e.g. totex models for gas distribution network companies at RIIO-GD2 and more recently for electricity distribution network companies).

In the context of residential retail cost assessment, given the extent of the historical data that will be available for cost assessment at PR24, we see a very weak case for econometric benchmarking models that *only* draw on companies' cost forecasts and do not use data on outturn costs. While costs and cost driver relationships may change over time, it would be better to allow this through the dynamic aspects of model specification, or in the adjustments used to move from modelled costs to allowances for the forthcoming price control period, than to omit historical data.

A more plausible approach is to use a dataset for the econometric modelling that includes both historical costs and forecast costs, as Ofgem has done.

One potential benefit of this type of approach, compared to the approach used by Ofwat for residential retail costs at PR19, is that it expands the sample size for the econometric modelling and *might* improve the accuracy of the estimated cost driver relationships in the model. However, we consider that this benefit is at risk of being overstated. And there are risks that the accuracy of the estimated cost driver relationships worsens by using forecast data, especially where different companies are influenced by different considerations and priorities when submitting their cost forecasts as part of their business plan.

We understand the argument that using a combination of historical cost data and forecasts of costs would help improve the reliability of econometric modelling, by increasing the sample size relative to modelling using historical data only. It is quite possible that some of the metrics of model performance that Ofgem and Ofwat use – such as the purported statistical significance of estimated coefficients – could be improved from the larger sample size. But these metrics can provide a misleading guide to the reliability of the benchmarking models given the potential for the cost forecasts to be unreliable and to distort the estimated relationships between retail costs and cost driver explanatory variables. The quality of data feeding into a model is an important consideration when deciding on what models or modelling approaches to use, further to any statistical results.

The relative benefits and risks arising from the use of forecast data will tend to vary across cases. For instance, for Ofgem's benchmarking of gas distribution companies at RIIO-GD2, Ofgem had a smaller sample of companies and short period of historical data than Ofwat will have for retail activities at PR24.

There is a further difference with Ofgem's regulation of energy network companies that seems relevant to retail cost assessment. The inclusion of cost forecasts in the input dataset can provide a way to take account of industry-wide changes over time that affect costs, beyond changes to cost driver variables and other factors such as input prices and ongoing productivity. For the energy network companies regulated by Ofgem, this feature might be useful in a context where (a) what energy network companies need to do is changing as part of decarbonisation and the energy transition; and (b) Ofgem's approach to cost assessment does not have an equivalent to the wide-ranging enhancement cost assessment arrangements used by Ofwat and so Ofgem's core benchmarking models may need to play a bigger role in allowing for increases in costs driven by changes to output and service requirements (where these are not captured by cost drivers used in the models). These issues of changes over time seem less relevant to retail cost assessment for water companies (though may be relevant to water companies' wholesale costs),

Finally, the use of company forecasts in the input data for benchmarking models is vulnerable to concerns that customers may not be protected against the risks of cost forecasts being too high. Despite the financial and reputational incentives arising from Ofwat's business plan assessment process, there are residual risks that company forecasts may over-estimate future levels of efficient cost, or lack sufficient ambition on the scale of further productivity improvements. In this context, there are significant risks to customers of allowances being set at too high a level if forecast cost data are routinely included in the modelling dataset. It may be possible to tackle these risks in the way that the benchmarking models are specified, and how model results are used to project cost benchmarks over the forthcoming price control period, but this would require considerable care and attention.

The PR19 approach, in which companies' cost forecasts are used to calculate an adjustment to be applied to modelled costs, allows for a degree of flexibility to mitigate these risks in a practical way. Ofwat seems more likely to apply such an adjustment if doing so provides for a more demanding set of allowances than other potential projection methods that use alternative sources of evidence. We doubt that Ofwat would apply an upward adjustment to the cost benchmarks derived from historical data simply because company forecasts exceed modelled costs. While this might represent a form of cherry-picking, it does not seem unreasonable in the circumstances. At the same time, as explained elsewhere, we consider it important that Ofwat takes steps to understand and mitigate risks that the company forecasts that are used to influence allowances are too low.

A4 Residential retail cost assessment at PR24: projection of cost benchmarks

Residential retail cost assessment at PR24: projection of cost benchmarks

Report for Bristol Water and Wessex Water – 12 April 2023

Reckon LLP

Data analysis | Economic regulation | Competition law

Table of contents

1: Introduction and summary 3
2: Conceptual foundation for the projection of cost benchmarks 9
3: Projections of cost benchmarks for the 2025-30 period 18
Appendix 1: Review of Ofwat's PR19 approach to residential retail cost assessment 33
Appendix 2: Specification of projection methods 47
Appendix 3: Supporting information for application of projection method 1 53

1: Introduction and summary

Background

As part of a project for Bristol Water and Wessex Water, we have developed and estimated a series of econometric benchmarking models of water companies' residential retail costs, over the historical period from 2013/14 to 2021/22. The approach to that work, and the main outputs of it, are set out in a separate report which we refer to in this document as our benchmarking report.[1]

In addition to the development and estimation of the benchmarking models, we have considered how to take results from those models and form projections of efficient cost benchmarks for each of Bristol Water and Wessex Water for the forthcoming price control period from 2025/26 to 2029/30. This has involved several phases of work:

- A detailed review of how Ofwat approached this aspect of its cost assessment for residential retail activities at PR19.
- Assessment of the main factors or phenomena to take into consideration when using results from econometric benchmarking analysis over a historical period as the basis for projecting cost benchmarks over a forthcoming price control period, and the identification and refinement of practical ways to do this.
- Developing projections of cost benchmarks for Bristol Water and Wessex Water for the 2025-30 period, drawing on some of the identified methods and approaches.

This report captures the outcome of these phases of work.

The projections, methods and evidence presented in this report are intended to help Bristol Water and Wessex Water as they develop their PR24 business plans, and as they respond to Ofwat consultations during the PR24 process. We have not sought to predict the cost benchmarks and approaches that Ofwat will use in its PR24 cost assessment.

Conceptual foundation for the projection of cost benchmarks

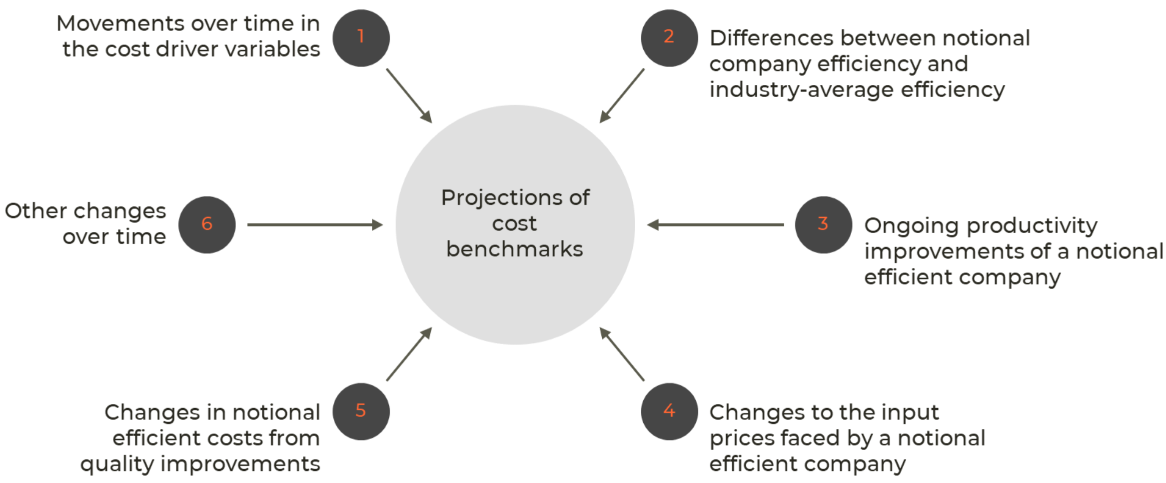
We adopt the established regulatory principle that price control cost allowance for a given company should be set by reference to the costs estimated for a notional efficient company, operating in that company's circumstances, rather than for the actual company in question. The notional efficient company might be defined as a company which operates at a greater level of cost efficiency than industry-average efficiency – though this need not be the case.

On this basis, there are a number of factors or phenomena to take into consideration when using results from econometric benchmarking analysis that covers a historical period as the basis for a cost benchmarking projected over a forthcoming price control period. Figure 1 presents a categorisation of these factors.

---

1 Reckon (2023) *Residential retail cost assessment at PR24: econometric benchmarking models,* report for Bristol Water and Wessex Water.

Figure 1 Summary of factors to reflect in projections of cost benchmarks



In section 2, we elaborate on the factors set out in Figure 1 and then describe how these factors might be taken into account in projecting cost benchmarks over the 2025-30 period.

Current regulatory practice in this area seems to have some methodological shortcomings. This is quite a complicated area, with plenty of pitfalls. There are significant risks of inconsistency between the different strands of analysis and assumptions that might be combined to produce projected benchmarks and, in turn, allowances. Furthermore, there are risks from focusing the assessment on only one type of methodology and source of evidence.

In section 2 we introduce five different "projection methods" that provide ways to take account of factors (1), (2), (3) and (4) from Figure 1 when projecting cost benchmarks over a forthcoming price control period. We summarise these projection methods in the table below (we provide a fuller description in appendix 2).

Table 1 Overview of alternative projection methods

Projection method	Brief introduction
1. Application of separate productivity and input price assumptions	This is what we see as the conventional approach in recent regulatory practice (e.g. applied by Ofwat and the CMA for wholesale base expenditure at PR19). It involves an adjustment for notional efficient levels of costs being applied to modelled costs over the forthcoming price control period, combined with separate adjustments for each of ongoing productivity improvements and input price effects for a notional efficient company.
2. Application of assumed unit cost trend	This has some similarities to (1) above except that rather than separate regulatory assumptions for input prices and ongoing productivity being determined and applied, a combined assumption on the trend in unit costs is determined which is intended to reflect the net effects of ongoing productivity improvements and changes in input prices for a notional efficient company. Ofwat used something close to this type of approach in setting allowances for operating expenditure at PR04 and PR09.

Projection method	Brief introduction
3. Extrapolation from econometric models that involve a time trend	This is an approach in which the effects of ongoing productivity and input price changes are captured by the inclusion of a time trend in the econometric models estimated on historical expenditure, and modelled costs for the forthcoming price control period are calculated by extrapolating that trend. This can be combined with an adjustment for notional efficient levels of costs.
4. Forward-looking adjustment based on business plan comparisons	This is based on the component of Ofwat's PR19 approach for residential retail costs which involved the calculation of an adjustment based on an upper quartile efficiency benchmark derived from comparisons of companies' business plan forecasts of retail costs over the 2020-25 period against modelled costs over that period. This adjustment is intended to take account of notional company efficiency, ongoing productivity and input prices in one go.
5. Business plan cost forecasts included in the input data for the benchmarking models	Under this approach, companies' business plan forecasts of retail costs would be included in the set of input data for the econometric benchmarking models. These forecasts would be expected to already incorporate companies' views on the impacts on costs of ongoing productivity and input price changes.

In addition, as we highlight in section 2, there are a number of further considerations which seem particularly important when making projections of benchmarks of residential retail costs. These concern the following:

- **Nominal versus CPIH-real projections**. If Ofwat retains the approach of no inflation indexation for the residential retail controls, as used at PR19 and envisaged in its PR24 final methodology, then the projected cost benchmarks should be calculated in nominal terms rather than in CPIH-adjusted prices.
- **Differences between nominal and CPIH-real average bills**. Ofwat's econometric models at PR19, and those we have developed, include an explanatory variable for the average size of bills in respect of bad debt related costs. The intuitive rationale is that, as the total amount billed increases, the level of bad debt related costs of an efficient retailer would also be expected to increase. The models provide strong evidence for this relationship. If nominal projections are needed, these should be calculated in a way that recognises the impacts on efficient residential retail costs of nominal (rather than simply CPIH-adjusted) bill changes over time.
- **Retail service quality interactions**. Depending on the projection method(s) used, and the evidence drawn on, there may be a case for adjustments to be made when calculating projected cost benchmarks for the 2025-30 period, to allow for expectations of retail service quality improvements over that period.

Projections of cost benchmarks for the 2025-30 period

We have applied three of the five projection methods summarised above: methods 1, 2 and 3. In each case, we have sought to take a proportionate approach within the time and resource constraints of the project and we have not sought to carry out the full scope of conceivable work under each method. We have not sought to implement projection methods 4 or 5 as part of the project. These rely on companies' business plan forecasts which are not yet available for PR24.

We describe our approach in section 3. In very broad terms:

- We consider projections calculated using two different suites of econometric models: a broader suite covering all 18 models presented as outputs in our benchmarking report and a narrower suite covering 10 of these 18 models.

- We calculate modelled costs over the 2025-30 period drawing on results from these econometric models and on a set of forecasts and working assumptions for cost driver variables.

- We draw on OBR forecasts for CPI (and in turn CPIH) inflation and, where needed, forecasts of changes over time in UK wage rates.

We summarise in Table 2 some of the key adjustments made under each projection method. No off-model adjustments for productivity, input price effects or cost trends are applied to modelled costs under method 3, as the modelled costs under method 3 already incorporate some allowance for these via the application of the time trend explanatory variables.

Table 2 Key adjustments under each projection method used to project from modelled costs

	Method 1: Application of separate productivity and input price assumptions	**Method 2:** Application of assumed unit cost trend	**Method 3:** Extrapolation from econometric models that involve a time trend
Upper quartile efficiency adjustment calculated from historical data	Reduction of around 9% applied to modelled costs in each year from 2025/26 to 2029/30		
Adjustment for assumed cost trend for notional efficient company up to 2030	N/A	Assumed cost trend of CPIH-2.0% per year	N/A
Adjustment for productivity growth of notional efficient company	Low assumption: 1% improvement per year High assumption: 2.9% improvement per year	N/A	N/A
Adjustment for impact on notional efficient costs of nominal input price changes	Adjustment which is on average 2.25% per year	N/A	N/A
Adjustment for impact on nominal costs of nominal vs CPIH-real changes in average bills	Adjustment which has effect of increasing costs by approx 0.8% per year	N/A	N/A
Uplift modelled costs by forecast CPIH up to 2029/30	N/A	Using OBR CPI forecasts to 2027/28 and assumption of 2.0% for 2028/29 and 2029/30 – implies average annual change of around 2.6% per year.	

Our approach to projection methods 2 and 3, and the high productivity scenario under method 1, reflect evidence on the unit cost reductions achieved by water companies in relation to residential retail activities in the period 2013/14 to 2021/22. These seem likely to reflect a phase of relatively high productivity improvements across the industry which resulted from a combination of some degree of historical inefficiency in the industry and the reforms to the regulatory approach to retail activities from PR14 onwards. We would not expect a productivity trend of this nature to be sustained over the long term. At the same time, it does not seem appropriate to base projections entirely on the 1% ongoing productivity trend that is often assumed in price control determinations and which we used for the low scenario under method 1.

At this stage in the PR24 process, we consider that the projection methods, and the approach and adjustments that we have used to apply them, are more important than the precise values calculated for the projected cost benchmarks.

Nonetheless, we have calculated a range of projected cost benchmarks for Bristol Water and Wessex Water, using the three projection methods and two model suites, over the 2025-30 period. We provide further details on the results and discussion of how they might be interpreted for the purposes of PR24 in section 3.

Potential updates to the cost benchmark projections

Given the very high levels of UK inflation at the time of preparing this report, and the particularly high uncertainty about inflation rates over the next few years, we suspect that some of the evidence and figures used in this report could become outdated quite quickly and may benefit from an update in light of emerging data on: water company costs; outturn inflation; updated OBR forecasts. In addition, as the PR24 process moves forward, companies will have updated forecasts for some of the cost driver variables used to make the projections and there may be value in updates to the projections to reflect these.

Interactions with uncertainty mechanisms

This report is concerned with methods and estimates for projecting residential retail cost benchmarks over the PR24 period. We do not consider in this report the case for price control adjustment mechanisms that would adjust allowances in light of outturn inflation (e.g. CPIH indexation of revenue allowances) or input price metrics (e.g. the labour RPE adjustment for wholesale controls at PR19). Nor do we consider here the type of industry-wide adjustment mechanisms involving outturn cost data that we have suggested in previous project work.[2] Depending on their design and scope, uncertainty mechanisms of this nature could reduce the importance of accurately reflecting factors summarised in Figure 1 when projecting benchmarks and setting ex ante allowances at PR24.

Structure of this report

The remainder of this report is organised as follows:

- Section 2 provides a conceptual foundation for the quantitative analysis that follows, setting out different methods that might be used to move from econometric benchmarking analysis carried out on historical data to make projections of cost benchmarks over the 2025-30 period, and highlighting a number of relevant considerations.

---

2 See section 5.6 of Reckon (2022) *The opportunities for a more coherent regulatory approach for Ofwat's funding of base expenditure and enhancements*.

- Section 3 presents our quantitative analysis to produce projections of cost benchmarks for Bristol Water and Wessex Water over the 2025-30 period, drawing on outputs from econometric benchmarking models and the application of the methods introduced in section 2.
- Appendix 1 provides a summary and review of a number of aspects of Ofwat's PR19 approach to residential retail cost assessment, with a particular emphasis on the adjustments for factors such as national company efficiency, ongoing productivity and input price effects.
- Appendix 2 provides further information on the methods that might be used to take the results from the econometric benchmarking models of residential retail costs and to make projections of cost benchmarks over a forthcoming price control period.
- Appendix 3 provides more detailed information on aspects of our approach to the application of projection method 1 as part of the analysis summarised in section 3.

2: Conceptual foundation for the projection of cost benchmarks

Introduction

In this section we consider ways in which we can draw on the results from econometric benchmarking models of residential retail costs to make projections of cost benchmarks over a forthcoming price control period. This section is organised as follows:

- Overview of factors to reflect in projections of retail cost benchmarks.
- Calculation of modelled costs over the forthcoming price control period.
- Alternative projection methods that might be applied to modelled costs.
- Allowing for the impact of nominal changes in average bills on efficient retail costs.
- The impact of service quality changes on notional efficient retail costs.
- Potential further adjustments for other changes over time.

In addition, we present in appendix 1 a summary and review of Ofwat's approach to projecting residential retail cost benchmarks. As explained in appendix 1, Ofwat's allowances for residential retail costs at PR19 were calculated using quite a novel approach with 50% weight attached to a forward-looking upper quartile efficiency adjustment, which was calculated using companies' business plan forecasts of retail costs. Aside from this, there seems to have been a mistake in Ofwat's approach to what the remaining 50% weight was applied to. We have drawn on insight gained from our review of the PR19 approach in developing this section.

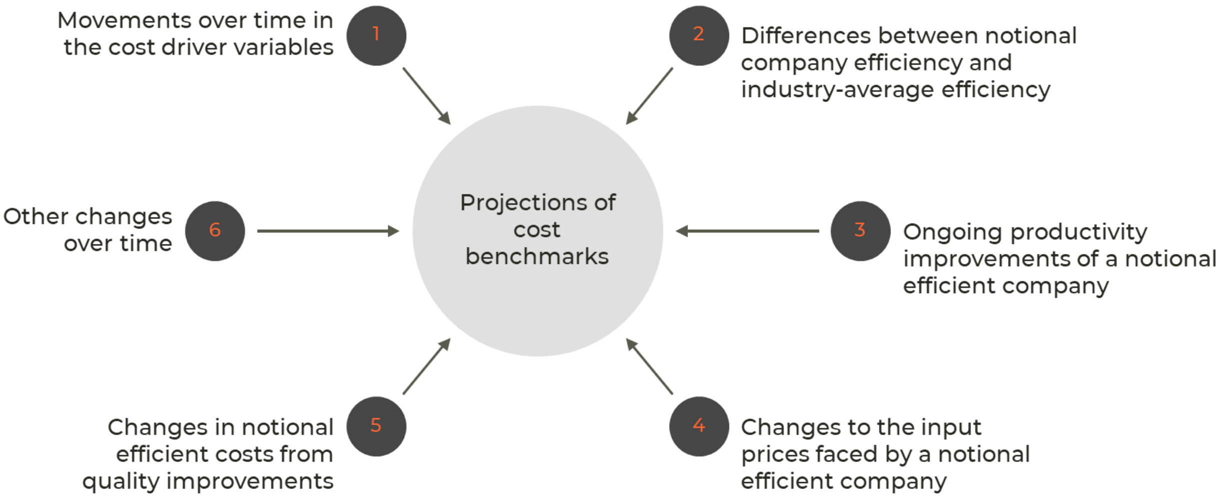
This section does not consider methodological issues relating to the development and choice of a suite of econometric benchmarking models, or the ways that results from different models might be triangulated.

Overview of factors to reflect in projections of retail cost benchmarks

We adopt the established regulatory principle that price control cost allowance for a given company should be set by reference to the costs estimated for a notional efficient company, operating in that company's circumstances, rather than for the actual company in question. The notional efficient company might be defined as a company which operates at a greater level of cost efficiency than industry-average efficiency – though this need not be the case (e.g. it might be defined as a company operating at an average or median level of efficiency).

On this basis, there are a number of factors or phenomena to take into consideration when using results from econometric benchmarking analysis that covers a historical period as the basis for cost benchmarks to be projected over a forthcoming price control period. Figure 2 presents a categorisation of these factors.

Figure 2 Summary of factors to reflect in projections of cost benchmarks



We summarise the six categories from Figure 2 as follows:

1. **Movements over time in cost driver variables**. Econometric benchmarking models estimated using historical data take account of a range of cost drivers for retail costs, through (a) the set of explanatory variables included in each model and (b), in the case of unit cost models, the denominator in the dependent variable (e.g. number of connected properties). For a given company, the values of these cost driver variables may change over time, between the historical period used for the econometric model estimation and the forthcoming price control period. When deriving estimated cost benchmarks over a forthcoming price control period from the econometric modelling results, some form of forecast or assumption for each of the cost driver variables over that future period is needed for each company.

2. **Differences between notional company efficiency and industry-average efficiency.** The modelled costs[3] derived from econometric benchmarking model will tend to produce cost benchmarks that reflect some concept of industry-average levels of efficiency (e.g. the average across companies and potentially also the average over the sample period if the models have a constant term rather than time trend or time dummy variables). The established regulatory practice of Ofwat, which was followed by the CMA in its wholesale cost assessment for the PR19 appeals,[4] is to set allowances for the forthcoming price control period in a way that reflects

---

[3] In this report, we follow Ofwat's convention in using the term modelled costs to refer to the cost predictions for a given water company (on a £ per household or £m basis) that are based on the predicted values from the econometric models, before the application of adjustments for factors such as upper quartile efficiency or productivity improvements over time. Modelled costs can be calculated for a historical period (using historical cost driver data) or a forecast period (using forecast cost driver data).

[4] The CMA said the following in the context of setting allowances for base expenditure for the wholesale price controls: *"Our cost models estimate how much it would cost the averagely efficient water company to cover base operations. However, we want to set cost allowances for a water company that is more than merely averagely efficient, and so we apply a 'catch up' efficiency challenge. Our decision is to use the company at the upper quartile as the benchmark and reduce the Disputing Companies' allowances accordingly. We consider this sets a challenging benchmark whilst*

a more demanding efficiency benchmark than the historical industry-average efficiency. This is sometimes described as a catch-up efficiency assumption. This factor does not represent catch-up from a company's *actual level of costs* to a notional efficient level of costs, but rather the catch-up that would be needed by a notional company operating at industry-average level of efficiency if it were to achieve some specified improved level of relative efficiency (e.g. perceived upper quartile efficiency).

3. **Ongoing productivity improvements of a notional efficient company**. It is reasonable to expect a notional efficient company to make ongoing improvements in productivity (or efficiency) over time. This is sometimes referred to as frontier-shift productivity or frontier-shift efficiency improvements. Ongoing improvements in productivity might take the form of cost reductions which enable (all else equal) a notional efficient company to reduce its costs in the forthcoming price control period relative to historical levels, or in the form of quality improvements which enable it to improve service quality relative to historical levels without increasing costs, or as a combination of cost reductions and service quality improvements.

4. **Changes to the input prices faced by a notional efficient company.** A notional efficient company's costs over the forthcoming price control period will be affected by changes over time in the input prices (e.g. wages and salaries and fees for inputs from third party suppliers). If allowances are needed for the notional efficient company's costs over the forthcoming price control period in CPIH-real terms, then the input price movements that are relevant here are movements in input prices relative to CPIH (sometimes referred to as real price effects or RPEs). If allowances are needed for the notional efficient company's costs over the forthcoming price control period in nominal terms (e.g. if there is no inflation indexation mechanism for the retail price control) then the input price movements that are relevant here are nominal movements in input prices.

5. **Changes in notional efficient costs from quality improvements**. Changes over time in the costs of a notional efficient company may reflect, in part, changes over time in the service quality provided by such a company. All else equal, the costs of a notional efficient company over the forthcoming price control period may increase relative to historical levels, at least insofar as there are incremental costs for an efficient company to achieve the expected improvements (though these might be offset to some extent by productivity improvements as highlighted under point (3) above).

6. **Other changes over time.** We define a residual category to capture any further factors which might lead to changes in the efficient costs of a company which are not properly captured in the cost drivers included in the suite of econometric models and not related to either productivity improvements or input prices. These might include, for example, external factors that affect retail costs (e.g. transitory impacts relating to the Covid-19 pandemic) which are not captured by econometric model specifications. It might also include changes to the relationship between

---

acknowledging the limitations of our econometric modelling (and the consequent risk that the company will have insufficient allowed revenue to ensure a base level of service". CMA (2021) *Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations: final report*; page 19.

efficient costs and a given cost driver included in a model, in cases where the model specification treats that relationship as constant over time.

Under the categorisation above, we treat input prices as separate from cost drivers, but we acknowledge that in other contexts it might be reasonable and useful to treat input prices as part of the set of cost drivers (and indeed to include input price explanatory variables in econometric model specifications). Our categorisation is intended to reflect the types of econometric models used at PR19 and which we expect to be used at PR24.

The remainder of this section considers ways in which projections of cost benchmarks might be calculated in a way that takes account of the factors (1) to (5) above. While adjustments to allow for factor (6) might be important in some cases, they are likely to be bespoke to the specific issue arising and we do not cover this in this section. Nonetheless, we give examples of the types of issues that might give rise to an adjustment under factor (6).

Calculation of modelled costs over the forthcoming price control period

Of the factors summarised above, the first is the most straightforward, at least at a methodological level. Ofwat's established approach is to produce forecasts for all explanatory variables over the forthcoming price control period, using company forecasts and/or its own forecasts.

While the details of any specific forecast might be disputed, the general principle of forecasting these variables seems well established and unproblematic. We do not consider the details of how forecasts are to be made in this report, other than for two points:

- First, there is a specific issue relating to the price base for the average bill explanatory variable which we discuss later in this section.
- Second, where econometric models include time trend variables or time dummy variables there are choices to be made about how to apply these when calculating modelled costs over a future period. The appropriate assumptions for time dummy variables, and how time trends should be applied, has interactions with the wider approach to making projections and the specific projection methods we discuss later in this section.

We provide more details on the second point in appendix 2 and we draw on both points for our quantitative analysis in section 3.

We use the terminology of "modelled costs over the 2025-30" period to refer to the predicted costs derived for each company by applying the estimated coefficients (from a suite of econometric models estimated using historical data) to forecasts or assumptions for the explanatory variables included in the suite of models over that period. Modelled costs for residential retail costs might be calculated on a cost per household basis or on an aggregate (£m) basis.

Alternative projection methods that might be applied to modelled costs

In this subsection, we summarise a number of alternative methods that provide ways to allow for factors (2), (3) and (4) when projecting cost benchmarks over a forthcoming price control period. To recap, these factors concern:

- differences between notional company efficiency and industry-average efficiency;

- ongoing productivity improvements of a notional efficient company; and
- changes to the input prices faced by a notional efficient company.

The choice of projection method can also have implications for the way that *modelled costs* over the 2025-30 period should be calculated, as highlighted above.

Current regulatory practice in this area seems to have some methodological shortcomings. This is quite a complicated area, with plenty of pitfalls. There are significant risks of inconsistency between the different strands of analysis and assumptions that might be combined to produce projected benchmarks and, in turn, allowances. Even the more conventional approach, which was applied by Ofwat and the CMA for PR19 for wholesale expenditure, seems to suffer from inconsistencies between the way that projections are made and the details of the econometric models used to calculate modelled costs.[5] Opportunities seem to be missed to draw on a wider, and potentially more relevant, set of data and evidence. Furthermore, there are concerns about the way that Ofwat used the business plan cost forecasts of a minority of water companies in setting residential retail cost allowances for all companies at PR19. See appendix 1 for further discussion of the approach of Ofwat and the CMA at PR19.

We have given this issue particular attention. We have sought to develop a number of coherent methods for projecting cost benchmarks that take account of the three factors above and avoid inconsistencies with the approach to calculating modelled costs. We refer to these as projection methods.

We describe and compare five different projection methods in appendix 2 and present an overview of these in Table 3. The application of these methods would need to be tailored according to whether projections of cost benchmarks are needed in nominal terms or relative to CPIH.

Table 3 Overview of alternative projection methods

Projection method	Brief introduction
1. Application of separate productivity and input price assumptions	This is what we see as the conventional approach in recent regulatory practice (e.g. applied by Ofwat and the CMA for wholesale base expenditure at PR19). It involves an adjustment for notional efficient levels of costs being applied to modelled costs over the forthcoming price control period, combined with separate adjustments for each of ongoing productivity improvements and input price effects for a notional efficient company.
2. Application of assumed unit cost trend	This has some similarities to (1) above except that rather than separate regulatory assumptions for input prices and ongoing productivity being determined and applied, a combined assumption on the trend in unit costs is determined which is intended to reflect the net effects of ongoing productivity improvements and changes in input prices for a notional efficient company. Ofwat used something close to this type of approach in setting allowances for operating expenditure at PR04 and PR09.

[5] For further discussion of this point, see the subsection on “The starting point for productivity and input price adjustments under Ofwat’s approach to wholesale controls” within appendix 1 of this report.

Projection method	Brief introduction
3. Extrapolation from econometric models that involve a time trend	This is an approach in which the effects of ongoing productivity and input price changes are captured by the inclusion of a time trend in the econometric models estimated on historical expenditure, and modelled costs for the forthcoming price control period are calculated by extrapolating that trend. This can be combined with an adjustment for notional efficient levels of costs.
4. Forward-looking adjustment based on business plan comparisons	This is based on the component of Ofwat's PR19 approach for residential retail costs which involved the calculation of an adjustment based on an upper quartile efficiency benchmark derived from comparisons of companies' business plan forecasts of retail costs over the 2020-25 period against modelled costs over that period. This adjustment is intended to take account of notional company efficiency, ongoing productivity and input prices in one go.
5. Business plan cost forecasts included in the input data for the benchmarking models	Under this approach, companies' business plan forecasts of retail costs would be included in the set of input data for the econometric benchmarking models. These forecasts would be expected to already incorporate companies' views on the impacts on costs of ongoing productivity and input price changes.

These five projection methods represent potential approaches to consider. Which of them are feasible and reasonable to apply in a specific case will depend on the nature of the evidence available and on a review of the results from their application. When setting allowances at a price control review, there may be merit in drawing on more than one projection method (e.g. by setting allowances using an average of results from different projection methods). Any single approach is likely to have benefits and drawbacks in the context in which it is applied.

Allowing for the impacts of nominal changes in average bills on efficient retail costs

One potentially important issue for residential retail cost assessment at PR24 is the way that an adjustment for changes over time in the levels of wholesale bills is factored into allowances for retail costs, especially given Ofwat's PR24 final methodology position that it will not allow CPIH indexation of the residential retail control.

Ofwat's econometric models at PR19 quite reasonably included an explanatory variable for the average size of bills in respect of bad debt related costs. The intuitive rationale is that, as the total amount billed increases, the level of bad debt related costs of an efficient retailer would also be expected to increase.

Based on updated data and further model development and review, we have seen support for models of bad debt costs that either (a) take the average size of bills as an explanatory factor as at PR19 or (b) have the effect of imposing an assumption of 1:1 on the relationship between average bill size and bad debt costs. In both cases, the calculation of modelled costs for the 2025-30 price control period for each company, on a £m basis, requires forecasts of each company's average bills over that period.

Care is needed on the price base used for the forecasts of average bills. If Ofwat retains the approach of no inflation indexation for the residential retail controls, as used at PR19 and envisaged in its PR24 final methodology, there is potential to systematically underfund water retailers if modelled costs for the 2025-30 period are calculated using forecasts of average bills on a CPIH-real basis (e.g. using forecasts of bills in 2022/23 prices) and no further adjustments are made to reflect the difference between nominal and CPIH-adjusted bills. At the same time, there is risk of double counting if the approach to retail cost assessment takes account of forecast changes in nominal average bills both through the forecast explanatory variables used to calculate modelled costs and through off-model adjustments for input prices or economy-wide inflation.

In this context, we see two main ways to incorporate expectations about each company's average bills over the 2025-30 period in the determination of allowances for residential retail costs in a scenario where – as Ofwat plans – there is no indexation of the residential retail controls:

- Use forecasts of average bills on a CPIH-real basis when calculating modelled costs for the 2025-30 period and combine this with off-model adjustments that can capture the impacts on efficient retail costs from the differences between nominal and CPIH-real average bills when determining nominal allowances for the 2025-30 period.

- Use forecasts of average bills on a nominal basis when calculating modelled costs for the 2025-30 period and then exclude the effects of nominal (rather than CPIH-adjusted) growth in average bills when applying adjustments (e.g. for productivity and input prices) to move from modelled costs to nominal allowances for the 2025-30 period.

As reflected in the approach we adopt in section 3, we see grounds to prefer the first option as it has benefits in terms of transparency.

There is some overlap between this issue relating to average bills and the broader consideration of allowances for movements in input prices over time. There may be scope for debate as to whether changes in average bills are to be treated as an input price issue. The more important thing is to adopt an internally coherent approach overall.

The impacts of service quality changes on notional efficient retail costs

One of the factors we identified earlier in this section concerned the effects on the costs of a notional efficient company of retail service quality improvements over time.

This report is not intended to provide a detailed consideration of the interactions between cost benchmarking and service quality. This is a complex matter. Reckon discussed this issue in the context of wholesale price controls for PR24 as part of a previous project.[6] But there are some important differences between wholesale and retail that have implications for these interactions.[7]

Nonetheless, we make some tentative comments below. For each of the five projection methods introduced above, we briefly comment on how the effects on the levels of efficient retail costs arising from service quality improvements over time might be considered in the residential retail cost assessment, and we highlight some examples of circumstances in which some form of adjustment

---

[6] See in particular section 2.3 and 2.4 of Reckon (2022) *The opportunities for a more coherent regulatory approach for Ofwat's funding of base expenditure and enhancements*.

[7] For example, for retail there are no separate allowances for enhancements which lie outside the main benchmarking models and the benchmarking models compare measures of costs (operating expenditure plus depreciation) rather than measures of expenditure used for wholesale cost assessment. Furthermore, the customer measure of experience (C-MeX) incentive is a type of relative incentive scheme across water companies with the median performing getting no incentive penalty or reward, without any explicit PCLs or incentive baselines being determined by Ofwat.

might be needed. In each case, the consideration of interactions between retail service quality and residential retail costs involves forming expectations about the levels of retail service quality (e.g. C-MeX scores), over the 2025-30 period, that would need to be achieved by a (notional efficient) company for it to face no financial penalty or reward on retail service quality.

Table 4 Interactions between retail service quality and projected cost benchmarks

Projection method	Tentative comments on service quality improvements
1. Separate productivity and input price assumptions	Provided that the expected retail service quality improvements (of a notional efficient retailer) over time are not too great, it may be reasonable to assume that those improvements are achieved by – and/or funded through – ongoing productivity improvements. If so, then the regulatory assumption on the annual rate of cost reduction that is assumed to be achieved via ongoing productivity improvements should be set in a consistent way (e.g. the more that ongoing productivity improvements is assumed to increase service quality the less scope there is for this to put downward pressure on costs, and vice versa) with the overall assumption on productivity growth reasonable in light of available evidence. In some cases there might be a case for an adjustment to retail cost allowances to fund the efficient costs of more substantial improvements in quality over time.
2. Application of assumed unit cost trend	The case for an adjustment will depend, in part, on how the unit cost trend has been set and what evidence it has drawn on. If the unit cost trend is based on historical trends in water industry retail costs, and if the *rate of improvement* in retail service quality (of a notional efficient retailer) is expected to be different in the forthcoming price control period compared to the past period used to assess historical trends, then there might be a case for an adjustment. If the unit cost trend is based on evidence from other sectors, it would be relevant, in principle, to consider what rate of quality improvements has been achieved by those sectors and how that compares with what is expected from a notional efficient water company in its provision of residential retail services – but this might be difficult to do in practice.
3. Extrapolation from econometric models that involve a time trend	Provided that the econometric model provides sensible results, the time trend in the model would be expected to reflect a range of factors that have acted on water company costs over the historical sample period, including productivity improvements, input price changes and any impact on efficient costs of making improvements in service quality over time. There might be a case for an adjustment if the expected *rate of improvement* in retail service quality over the forthcoming price control period is significantly different to the rate of improvement experienced over the historical period used to estimate the time trend.
4. Forward-looking adjustment based on business plan comparisons	In preparing their business plan cost forecasts for residential retail activities, companies may make explicit or implicit assumptions about the levels of retail service quality that they will need to achieve to avoid financial penalties under the C-MeX incentive scheme. We have not identified a role for a separate adjustment for service quality to be applied under this projection method, but the potential for companies' explicit or implicit assumptions around retail service quality to be inaccurate is a factor that affects the accuracy of this project method.
5. Business plan cost forecasts included in the input data for the benchmarking models	See for point (4) above.

The comments in the table above are not intended to be comprehensive of all potentially relevant considerations (e.g. we do not cover issues that might arise from increasing marginal costs of service quality improvements).

Potential further adjustments for other changes over time

The factors that we have discussed in the subsections above are of general relevance and applicability. In addition, there may be some further case-specific factors that mean that levels of efficient residential retail costs will be different in the future compared to the past.

We provide some examples of the types of issues that may arise:

- **Transitory factors affecting efficient levels of costs not captured in modelling**. The efficient levels of retail costs could differ in the future compared to the past due to transitory factors (historically or in the forthcoming price control period) that are not captured through the econometric model specifications and the calculations of modelled costs. The impact of Covid-19 on retail costs is one possible example.

- **Changes in cost driver relationships not captured in models**. Ofwat's econometric models for benchmarking costs at the retail and wholesale levels assume that the relationship between costs and each cost driver is constant over time (e.g. the impact on costs of an X% increase in a specific cost driver variable is the same over the historical sample period) and Ofwat generally sets cost allowances under an assumption that the same relationship applies in the future. In some cases, there may be an argument that certain factors mean that the relationship will change over time. For instance, changes in water companies' tariff structures and social tariff arrangements, which have the effect of increasing the scale of bill reductions offered to lower-income households, might reduce the extent to which companies in more deprived areas face higher bad debt costs – implying changes over time in the coefficient for explanatory variables linking deprivation to retail costs.

- **Limitations in forecasts of explanatory variables**. It is also possible that the forecasts available to be used for explanatory variables over the forthcoming price control period do not properly capture changes over time affecting those variables (e.g. if forecasts are based on extrapolation). In general, it would be better to address this by changing the forecasts of explanatory variables to reflect expected future conditions, but in some instances this might not be done and there may be a case for an adjustment to modelled costs before setting allowances.

For some of these factors, it might be possible to take appropriate account of them through the use of alternative specifications for the econometric models (e.g. models that allow the coefficient on a specific cost driver to differ across different years or parts of the sample period). This may be difficult to achieve in some cases. Instead, there could be a role for adjustments applied to some or all company allowances if there is good evidence to think that cost driver relationships could differ over the forthcoming price control period compared to what is reflected in modelled costs.

3: Projections of cost benchmarks for the 2025-30 period

Introduction

In this section we present quantitative analysis which uses currently available data to combine results from econometric benchmarking of companies' historical costs for residential retail activities with a series of assumptions and adjustments to form projections of efficient cost benchmarks for the 2025-30 period for Bristol Water and Wessex Water.

We apply three of the five projection methods presented in section 2, which offer different ways to allow for movements over time in cost driver variables, notional company efficiency, ongoing productivity improvements, and input price changes. These are:

- Projection method 1: Application of separate productivity and input price assumptions.
- Projection method 2: Application of assumed unit cost trend.
- Projection method 3: Extrapolation from econometric models that involve a time trend.

In each case, we have sought to take a proportionate approach within the time and resource constraints of the project and we have not sought to carry out the full scope of conceivable work under each projection method.

We have not sought to implement projection methods 4 or 5 as part of the project. These rely heavily on companies' business plan forecasts for the forthcoming price control period, which are not available for PR24 at this stage.

We have made projections on a cost per household basis, which is more robust to the uncertainty over forecasts of customer numbers over the 2025-30 period than projections made on an aggregate (i.e. £m) basis.

At PR14 and PR19, Ofwat set residential retail price controls that did not involve CPIH indexation of revenue limits. In its PR24 final methodology, Ofwat said that it intended to maintain this aspect of its approach. On this basis, Ofwat's cost assessment process for residential retail activities at PR24 will determine cost allowances that will apply in nominal terms over the 2025-30 period, rather than determining cost allowances on a CPIH-adjusted basis (e.g. in 2022-23 prices). In this context, we consider that the most relevant basis for presenting projections of residential retail cost benchmarks is on a nominal basis, rather than on a CPIH-adjusted basis.

The remainder of this section takes the following topics in turn:

- Selection of econometric models and approach to triangulation.
- Forecast explanatory variables and calculation of modelled costs.
- Forecasts of CPIH inflation.
- Notional efficient company catch-up adjustment.
- Approach to the application of projection method 1.

- Approach to the application of projection method 2.
- Approach to the application of projection method 3.
- Residential retail cost benchmarks calculated under the three projection methods.
- Discussion of projected benchmarks and implications for PR24.

Selection of econometric models and approach to triangulation

Our work to develop econometric benchmarking models of residential retail costs is set out in a separate report prepared for Bristol Water and Wessex Water, which we refer to as our benchmarking report.

We have used two different suites of models for our projections:

- **Broader suite**. This suite contains all 18 models from the set presented in our benchmarking report.
- **Narrower suite.** This is a narrower suite of 10 models that excludes some models from the broader suite. Specifically it only includes models that have a time trend variable, and those models of other retail costs for which the dependent variable is on a cost per household basis rather than on a cost per service basis. This narrower suite of models was submitted to Ofwat by Wessex Water in its model submission to Ofwat in January 2023.

For each suite, we give the applicable models equal weight within the cost category to which they apply when calculating modelled costs. We then aggregated the modelled costs (in £m) from the two categories that we modelled separately, bad debt related costs and other retail costs, and expressed the results on a £ per household basis.

Forecast explanatory variables and calculation of modelled costs

For the purposes of our projections, we need forecasts or assumptions for the explanatory variables (and the denominator in the dependent variable in some cases) used in the suites of econometric models, over the 2025-30 period.

As a starting point, we sought to use forecasts from the companies concerned where these were available. This was as follows:

- Bristol Water provided its latest internal forecasts for the 2025-30 period relating to the number of households supplied and meter penetration.
- Wessex Water was not able to provide internal forecasts at this stage.

For the explanatory variable relating to the average size of residential retail bills (residential revenue divided by number of households supplied) we made the assumption, in agreement with Bristol Water and Wessex Water, that average bills will change in line with our forecasts of CPIH since their observed levels in 2021/22.

For the purposes of model input data, we used average bill figures in 2017/18 prices (2017/18 prices was the price base for the cost input data that we used). This meant that, combined with our assumptions of average bill growing in line with CPIH, the data on average bills for each year over the 2025-30 period that we used to calculate modelled costs was the same value as the model input data for 2021/22.

For other cost driver variables not covered above we took the following approach:

- If there was a forecast of that variable up to 2024/25 from the set of data files published with Ofwat's PR19 FD, we took that value for 2024/25 and applied it over the period to 2029/30.
- If there was no such forecast from the PR19 FD we took the latest available value from our historical dataset (i.e. 2021/22) and applied this over the period to 2029/30. This was the case for the deprivation and arrears risk variables as well as the mix of retail services between water and wastewater.

The assumptions on cost driver variable forecasts that we used for the purposes of this report reflect the information available at present and a proportionate approach to the analysis. We note the following:

- As the PR24 process moves forward, companies will have updated forecasts for a number of the cost driver variables and the projections could be updated to reflect these.
- There could be a role for more detailed work to consider how the ONS deprivation and Equifax credit risk variables are likely to evolve over the period to 2030, which could feed into updated or refined projections.

Further to the points above concerning cost driver variables, there is a need to tailor the way that we calculate modelled costs depending on which projection method the modelled costs is to be applied to. For instance, projection method 3 requires the use of econometric benchmarking models that involve a time trend; this means that only a subset of the models from the broader suite can be used for this projection method. We specify below where there are differences in modelled costs between projection methods. In addition, we specify how we have applied time dummy variables and time trends when calculating modelled costs for the 2025-30 period.

Table 5 Further details on assumptions used for the calculation of modelled costs

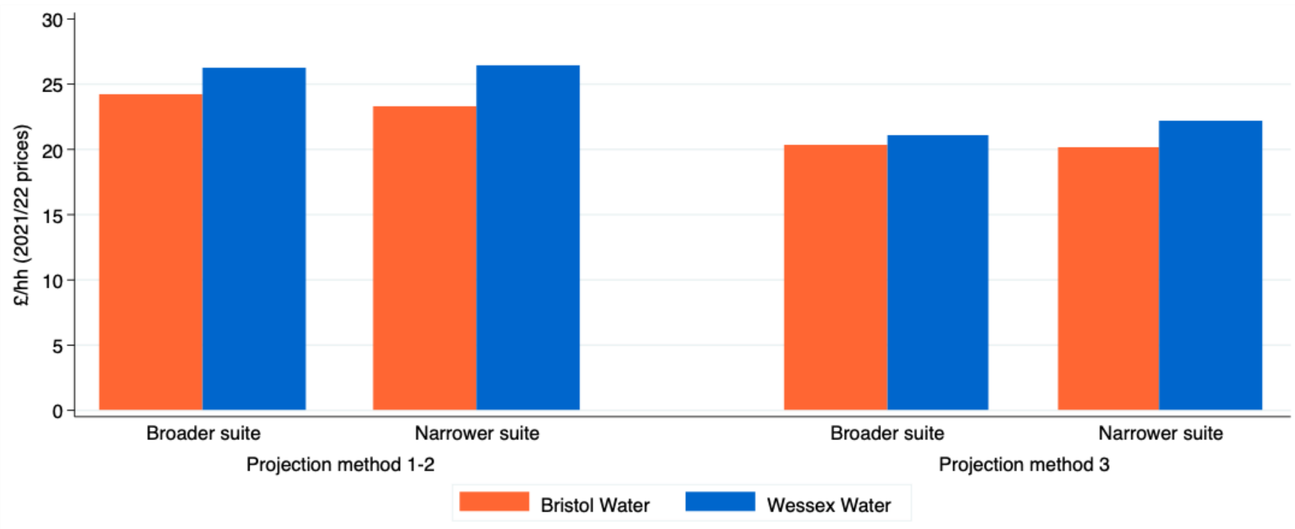
Metric	Projection method 1	Projection method 2	Projection method 3
Restrictions to suite of models used	None		Only use models with a time trend
Application of time trend in the case of models with a time trend variable	Time trend NOT applied beyond 2021/22 (this was done by fixing the financial year to 2021/22 for the purposes of the time trend)		Time trend applied up to 2029/30
Application of time dummies in the case of models that do not have a time trend variable	Time dummy for 2021/22 set to one; all others time dummies set to zero		N/A
Application of Covid dummies in the case of bad debt models that have a time trend and Covid dummies	All three Covid dummies set to zero		

We chose assumptions for the time trend and time dummy elements under projection methods 1 and 2 with a view to deriving modelled costs for the 2025-30 period that represent costs for a company operating at a level of productivity, and facing input prices, reflecting those of a notional efficient company in 2021/22. We then use subsequent adjustments, outside the calculation of modelled costs, to allow for productivity and input price changes in the period to 2030. In contrast, under method 3, the time trend used in the calculation of model costs is intended to make an allowance for input price changes (relative to CPIH) and notional efficiency company ongoing productivity improvements within the calculation of modelled costs.

Given that the model input data on costs, and our assumptions on average bills in the 2025-30 period, were in 2017/18 prices, modelled costs calculated on the basis set out above would also represent cost benchmarks for each company in 2017/18 prices. As a final step, we updated the price base for modelled costs, from 2017/18 prices to 2021/22 prices using the same CPIH series that we use to convert costs in nominal terms to costs in 2017/18 prices.

On the basis set out above, we present the modelled costs for Bristol Water and Wessex Water below, for the two different suites of models and, as applicable, under projection methods 1, 2 and 3. Since modelled costs are calculated on the same basis for methods 1 and 2 we have grouped these. The figures in the chart are the annual average for the 2025-30 period on a cost per household basis in 2021/22 prices.

Figure 3 Modelled costs: annual average 2025-30 (2021/22 prices)



The modelled costs for projection method 3 are systematically lower than the corresponding modelled costs for methods 1 and 2. This reflects the application of the time trend variables in the calculation of modelled costs under method 3, and not under methods 1 and 2, in a context where the estimated coefficients on the time trend variable imply significant reductions in costs over time, relative to CPIH.

Forecasts of CPIH inflation

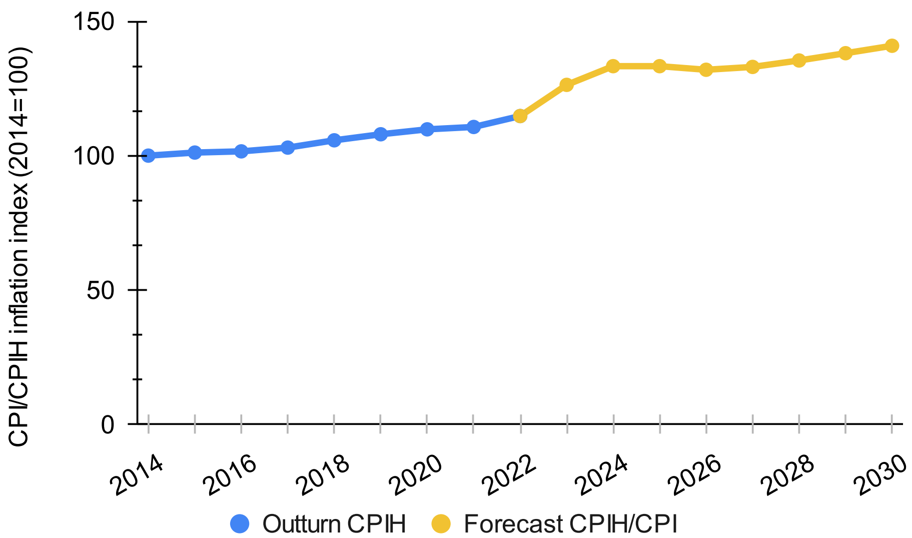
For all three of the projection methods that we use, we use forecasts of CPIH inflation for the period from 2021/22 to 2029/30. For projection method 1, we use CPIH forecasts in calculating an approximate nominal input price index for a water retailer and to take account of the impacts on nominal retail costs arising from nominal increases in average bills that are not captured in modelled costs. For methods 2 and 3, we use CPIH forecasts to convert figures in 2021/22 price base to figures in a nominal price base.

Our approach to CPIH inflation forecasts has been as follows:

- Consistent with the way in which Ofwat's historical cost data is deflated from nominal to real prices, we use measures of the annual change in CPIH (or CPI) which represent a measure of the change in the average value of CPIH/CPH in one financial year (e.g. monthly values of index from April to March) relative to the corresponding average value of CPIH/CPH in the previous financial year.
- Consistent with recent regulatory practice, we use OBR forecasts of CPI inflation as the primary source of forecasts for CPIH inflation.
- We took OBR forecasts from its *November 2022 Economic and fiscal outlook*.
- The OBR forecasts for CPI are made up to 2027/28. For the last two years of our forecast period, we assumed CPI and CPIH growth of 2% per year. This represents a return to inflation rates around the Government's CPI target. This seemed a reasonable assumption, especially given that the OBR's forecasts involve a period of relatively high and then relatively low inflation over the next few years, before a more normal rate of 1.8% in 2027/28.

Figure 4 shows an index covering CPIH inflation in the period of our historical dataset used for the econometric modelling, and the forecast period up to 2029/30. Under the OBR forecasts, the current period of relatively high UK inflation involves a peak in inflation rates in 2022/23 (10.1%) after which the rate slows and becomes very low and even negative (-1.0% in 2025/26). On the basis of the OBR forecasts, the 2025-30 price control period will be one of relatively low economy-wide inflation. However, there is considerable uncertainty about future UK inflation.

Figure 4 Outturn and forecast CPIH inflation



Notional efficient company catch-up adjustment

The calculation of modelled costs using forecasts of explanatory variables, as described earlier in this section, can be seen to produce cost benchmarks for an averagely efficient company.

For each of the three projection methods that we have used, there is a role for an adjustment which is intended to adjust for differences in costs between those of an averagely efficient company and those of a notional efficient company.

For the purposes of this project, we have sought to maintain consistency with Ofwat and CMA precedent on the use of such an adjustment. We have used the upper quartile of the triangulated modelled costs. An upper quartile benchmark company was used by Ofwat for residential retail at PR19 and by the CMA for wholesale base costs at PR19. This aspect of current regulatory practice might be debated further (e.g. on asymmetric risk grounds) but seems firmly entrenched in Ofwat and CMA practice and it was not the role of this project to explore this matter further.

Following the approach taken at the wholesale level by Ofwat and the CMA, we calculated efficiency ratios for each company by aggregating the last five years of outturn costs and comparing these against modelled costs over that historical period, and then calculated an adjustment factor as one minus the efficiency ratio of the company identified as upper quartile in the ranking of efficiency ratios.

On this basis, we calculated a reduction of around 9% to be applied to modelled costs in each year for the purposes of the notional efficient company catch-up adjustment. The precise value of the adjustment varied slightly between the broader model suite (9.34%) and narrower model suite (8.96%). We applied the same adjustment under all three projection methods.

Approach to the application of projection method 1

We now provide some further information on our approach to the application of projection method 1. Under this method we apply three different types of adjustments, further to the notional efficient company catch-up adjustment, to derive projected cost benchmarks for the 2025-30 period in nominal terms. These are adjustments for the following:

- **Ongoing productivity**: the productivity growth of a notional efficient company, in relation to residential retail costs, over the period from 2021/22 to 2029/30.
- **Nominal input prices**: changes in the nominal input prices faced by a notional efficient company, in relation to residential retail costs, over the period from 2021/22 to 2029/30.
- **Nominal bill inflation**: the impact on differences between nominal and CPIH-adjusted average bills on the costs of a notional efficient company.

We summarise our approach to each of these adjustments in the table below. We provide some further context and information on our approach to projection method 1 in appendix 3. The approach we adopted, and the figures arising from it, are very approximate and there would be considerable scope to refine the approach. However, for the purposes of this report, given our use of two other projection methods besides method 1 and the inherent uncertainty in the overall exercise, we did not consider it appropriate to focus our time on a more sophisticated application of method 1.

Table 6 Summary of approach and assumptions specific to projection method 1

Adjustment type	Summary of approach to the adjustment
Ongoing productivity We used two different assumptions which were based on different approaches and sources	**Lower productivity scenario** We used an assumption of 1% improvement in productivity per year. This figure has substantial regulatory precedent and, in turn, draws on EU KLEMS data on productivity growth for the UK economy and sectors within it. There is an argument that, in some contexts the 1% assumption might be seen as a central or high scenario for productivity growth, but given the indications of higher productivity growth amongst water retailers since PR14, we treat it as a low scenario.
	Higher productivity scenario We made a very approximate estimate of the historical annual average productivity growth of a notional efficient water retailer over the sample period for our econometric modelling. This involved taking our estimate of the average annual change in unit costs between 2013/2014 and 2021/22 for a notional upper quartile retailer (-2.0% per year relative to CPIH) and adjusting for a high-level approximation for the average annual change in input prices relative to CPIH faced by such a retailer over this period. Our approach for the average annual change in unit costs for a notional upper quartile retailer is described in more detail later in this section when we elaborate on projection method 2. Our approach for the changes over time in input prices for a notional upper quartile retailer involved the creation of an index of input price growth (relative to CPIH) and is described in more detail in appendix 3. The implied annual average change in input prices (relative to CPIH) under this approximate approach was 0.9% per year. Based on this approach, we used a figure of 2.9% per year for the ongoing productivity improvement of the notional efficient company under the higher productivity scenario. This can be seen as a rate of productivity improvement that is consistent with a cost trend of CPIH-2% per year and a CPIH-adjusted input price trend of 0.9% per year.
Nominal input prices	We created an index of the approximate nominal input prices that would be faced by an efficient residential retailer over the period from 2021/22 to 2029/30, drawing in particular on OBR forecasts of CPI inflation and changes in wages and salaries, and a decomposition of the costs of a notional water retailer into three main categories: bad debt costs; labour costs; residual costs. We provide more details on this approach in appendix 3. We used this index to calculate a series of annual adjustment factors for the period from 2021/22 to 2029/30. The implied annual average change in nominal input prices under this index was 2.25%.
Nominal vs CPIH-adjusted average bill forecasts	We created an adjustment factor for differences between forecast nominal and CPIH-adjusted bill increases, to take account of our objective to produce cost projections in nominal terms in a context where the average bill forecasts used in the calculation of modelled costs is on a CPIH-adjusted basis (2021/22 prices). We did this as follows. • We set the adjustment factor to 1 in 2021/22. • For each subsequent year we applied an adjustment which was calculated as 0.976 * the assumed share of bad debt costs (using the same ratio for our input price index above) * our forecast CPIH change for that year. The figure 0.976 is the average of the estimated coefficient on the average bill variable in our models of bad debt related costs per household, and can be interpreted as implying that a 1% increase in average bills would increase bad debt related costs by 0.976%

Approach to the application of projection method 2

A benefit of projection method 2, compared to the types of approach used for wholesale at PR19 (i.e. variants on method 1), is that data on changes over time in water companies' residential retail costs can be brought to bear on the analysis in quite a straightforward way. Historical trends in water retailer unit costs are clearly relevant to the projection of future trends in retail cost benchmarks.

In addition to using data on water industry retail costs, there are other possible sources of data on unit cost trends (or similar) that might be considered under projection method 2. For instance, it is possible to calculate trends in unit costs from the EU KLEMS dataset, so that this data source is used to inform on the combined impact of ongoing productivity and input prices in various UK sectors, rather than the conventional role for it to inform on ongoing productivity only.[8] However, it is questionable what additional insight this would add, given that the data would be for broad UK sectors that are clearly different in many ways from retail activities in the water industry and data from the EU KLEMS dataset are often drawn on under projection method 1. Similarly, there may be other regulated industries for which a unit cost trend could be calculated, but it is difficult to see what evidential power this would have in the context of PR24, given differences between regulated industries.

Given these issues, for the purposes of our projections we focused on the application of projection method 2 using data on water industry retailers, for which we consider the case for using method 2 is strongest.

Table 7 summarises several different types of analysis of water companies' residential retail costs which we carried out, which are reflected in metrics (A), (B) and (C) in the table. For each metric, we show figures for cost trends over the full 9-year period of our historical data period and for the last five years of that period. Two figures are presented in each cell for metrics (B) and (C) and these represent figures for the broader suite of econometric models and the narrower suite of econometric models respectively.

Table 7 Metrics feeding into assumptions on unit cost benchmark

Metric	2013/14 to 2021/22	2017/18 to 2021/22
A. Annualised rate of change (compound annual growth rate) in the total residential retail costs across the industry on a £ per household basis.	-1.8% per year	-1.8%
B. Changes over time in the triangulated unit cost benchmark derived from the econometric benchmarking models for a hypothetical company for which the cost driver variables (e.g. average bill size and meter penetration) are assumed to be constant over the historical period.	-1.9% / -1.9%	-1.5% / -1.2%

[8] See for example figure 17 from Reckon (2011) *Productivity and unit cost change in UK regulated network industries and other UK sectors: initial analysis for Network Rail's periodic review.*

Metric	2013/14 to 2021/22	2017/18 to 2021/22
C. Annualised changes in the unit cost benchmark, derived from our econometric benchmarking models, for a hypothetical company (a) that operates at an upper quartile level of efficiency in every year; and (b) for which the cost driver variables are constant over the historical period. Taking each year in turn, we calculated the ratio of each company's outturn total retail costs in that year to their triangulated modelled total retail costs and then calculated the upper quartile value across the distribution of those ratios. We then applied a dynamic annual upper quartile adjustment to the triangulated unit cost benchmark used for metric (B) above.	-2.0% / -2.0%	-2.5% / -2.1%

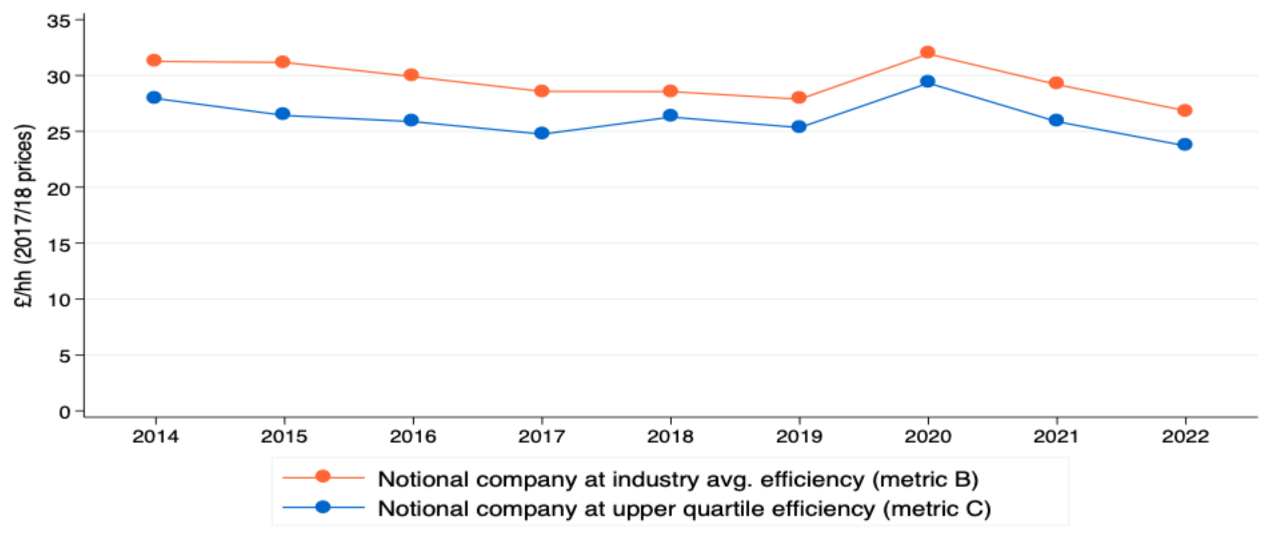
Another possible type of analysis to inform on the unit cost trend in residential retail costs per household is the time trend estimated from those econometric benchmarking models in our suites which involve a time trend. However, this is something that we use directly in projecting cost benchmarks under projection method 3, so we have not sought to duplicate it here. That said, estimates from metrics (B) and (C) will reflect, in part, the time trends from such models.

Metric (A) above is easiest to calculate but has the potential to be misleading if changes in unit costs have been driven by factors besides productivity and input price changes (e.g. changes in unit costs driven by changes over time in cost drivers relating to meter penetration, average bills or economic deprivation). The other two metrics largely avoid this problem because they draw on results from the econometric benchmarking models in a way that control for changes over time in cost driver variables.

We consider that metric (C) is the most relevant and internally consistent to use. This metric is directed at the cost trend for a notional efficient company – rather than the cost trend for a company operating at an industry-average level of efficiency (which is an implicit feature of the estimated annual changes under the other two metrics).

We present in the chart below a profile of unit costs over time for metrics (B) and (C).

Figure 5 Estimated profile over time for unit costs of average-efficient and UQ-efficient company



While the cost benchmark for an upper quartile company under metric (C) are lower in each year than for the cost benchmark for an average-efficient company under metric (B), the overall change between the start and end of the sample period is very similar in percentage terms. This chart indicates that the relatively large decreases in unit costs observed across the industry are not simply due to worse-performing companies catching up to other companies' levels of efficiency.

From Figure 5, we can see clear effects from the Covid-19 pandemic in metrics (B) and (C) in 2019/20 and also the potential impacts in 2020/21 and 2022/23. When 2022/23 data is available, it would be useful to extend the analysis to see how this affects the calculation of changes over time in unit costs (although the costs reported in 2022/23 will also be affected by the recent period of abnormally high inflation).

For the purposes of projection method 2, our approach has been to:

- Use an approximate assumption of a cost trend of CPIH-2% per year, based on the analysis set out above (over the full sample period) and in particular our estimates of the modelled unit cost reductions for a notional efficient upper quartile company.
- Recognise that this assumed rate of change in unit costs is likely to reflect a period of high productivity growth arising from the PR14 reforms (e.g. these reforms included separate price controls targeted on residential retail activities, emphasis on benchmarking to set allowances and companies bearing more financial exposure to their bad debt costs), which may not be sustainable over the longer term.

Applying this assumed cost trend of CPIH-2% to the modelled costs that we calculated for the 2025-30 period provides cost benchmarks for that period in 2021/22 prices. To produce projections in nominal terms, we apply the forecasts of CPIH over the period to 2030 as described earlier in this section.

Approach to the application of projection method 3

Under projection method 3, we calculate modelled costs for the 2025-30 period in a way that incorporates some allowance for changes in productivity and input price changes (relative to CPIH) in the period up to 2030, via the time trend variables used in the econometric models.

The estimated time trend coefficients vary across the models and differ in particular between the models of bad debt costs and models of other retail costs. For instance, the time trend coefficients in the models of bad debt costs per household imply unit cost reductions of around 4.5% per year and the time trend coefficients in models of other retail costs per household imply unit cost reductions of around 2% per year. In turn, the impact of the time trend on modelled total residential retail costs per household differs between companies because the mix of modelled bad debt costs and modelled other costs varies between companies.

Across companies, the effect on modelled costs of applying the time trend over the period 2021/22 to 2029/30 is significantly greater on a triangulated basis across the models used than the assumed trends in unit costs of CPIH-2% used for projection method 2.

The time trend estimated over the historical sample period is not necessarily a good guide for the future. In particular, we recognise that the estimated trend in unit costs is likely to reflect a period of high productivity growth arising from the PR14 reforms, which may not be sustainable over the longer term.

The modelled costs calculated under projection method 3 are in 2021/22 prices. To produce projections in nominal terms, we apply the forecasts of CPIH over the period to 2030 which we used for the purposes of projection method 1, which are largely based on OBR forecasts of CPIH.

Residential retail cost benchmarks calculated under the three projection methods

We now turn to present results from the application of the three projection methods outlined above.

Before doing so, we summarise the key assumptions and adjustments used under each projection method in Table 8. This table shows where adjustments are common across methods and how the methods differ in terms of the types of adjustments applied. The figures in the table do not represent the specific assumptions and adjustment that we applied to make nominal projections of the cost benchmarks to 2025-30 (e.g. the figures we used were on an annual basis and calculated more precisely). But they provide an indication of the magnitude of the adjustments. We apply the adjustments in 2022/23 and each subsequent year, reflecting our approach of calculating modelled costs in the 2021/22 price base.

Table 8 Key adjustments under each projection method used to project from modelled costs

	Method 1: Application of separate productivity and input price assumptions	**Method 2:** Application of assumed unit cost trend	**Method 3:** Extrapolation from econometric models that involve a time trend
Upper quartile efficiency adjustment calculated from historical data	Reduction of around 9% applied to modelled costs in each year from 2025/26 to 2029/30		
Adjustment for assumed cost trend for notional efficient company up to 2030	N/A	Assumed cost trend of CPIH-2.0% per year	N/A
Adjustment for productivity growth of notional efficient company	Low assumption: 1% improvement per year High assumption: 2.9% improvement per year	N/A	N/A
Adjustment for impact on notional efficient costs of nominal input price changes	Adjustment which is on average 2.25% per year	N/A	N/A
Adjustment for impact on nominal costs of nominal vs CPIH-real changes in average bills	Adjustment which has effect of increasing costs by approx 0.8% per year	N/A	N/A
Uplift modelled costs by forecast CPIH up to 2029/30	N/A	Using OBR CPI forecasts to 2027/28 then assume 2% change - effect around 2.6% on average per year from 2021/22	

The table helps show some relevant differences between the projection methods:

- No off-model adjustments for productivity, input price effects or cost trends are applied to modelled costs under method 3, as the modelled costs under method 3 already incorporate some allowance for these via application of the time trend explanatory variables.
- In projection method 1 there are explicit assumptions for ongoing productivity and input price changes, whereas the combined effects of these are implicit with the cost trends applied under methods 2 and 3.
- No uplift for CPIH inflation is needed as a final step in method 1, because this method involves an explicit adjustment for nominal input price inflation and an adjustment to take account of the increases to nominal costs arising from our assumption that average bills will grow from 2021/22 onwards in line with the CPIH.

As explained earlier in this section, we have used two different model suites when triangulating results from the econometric benchmarking models. We have three projection methods and, for the first of these, two sets of projections which vary according to whether our high or low assumption on ongoing productivity growth is applied. Taken together this means that, for each of Bristol Water and Wessex Water, we have eight different projections for the efficient residential retail cost benchmark for each year in the 2025-30 period.

We provide charts for each of Bristol Water and Wessex Water which show the eight projections for cost benchmarks over the 2025-30 period. To limit complexity in the charts we present an annual average across the five-year period, although there are differences across years.

Figure 6 Projected efficient cost benchmarks for Bristol Water (nominal)

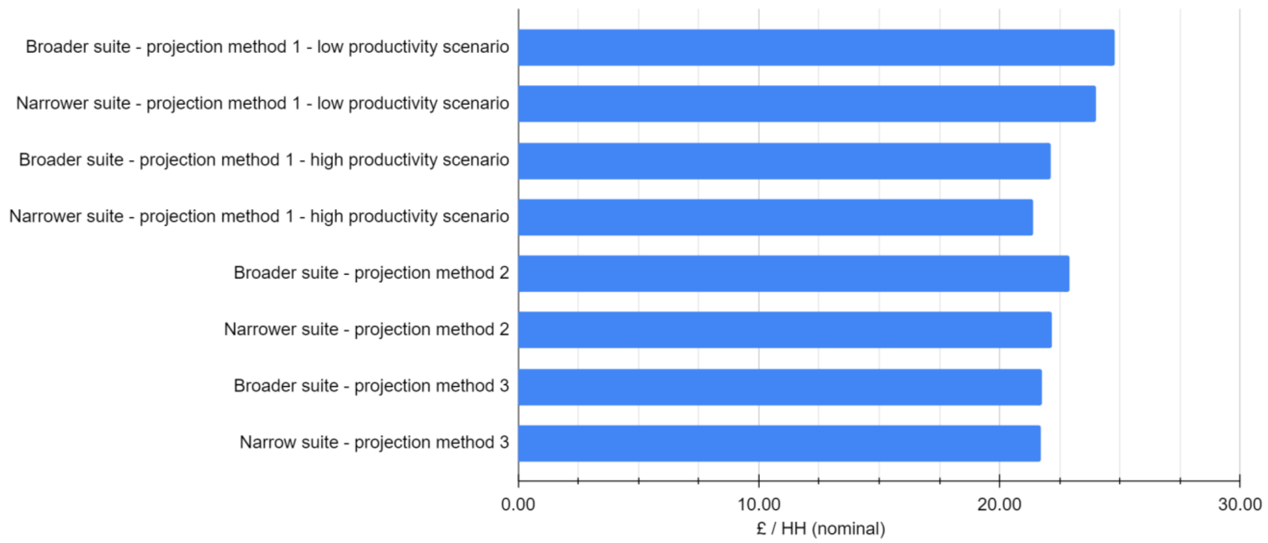
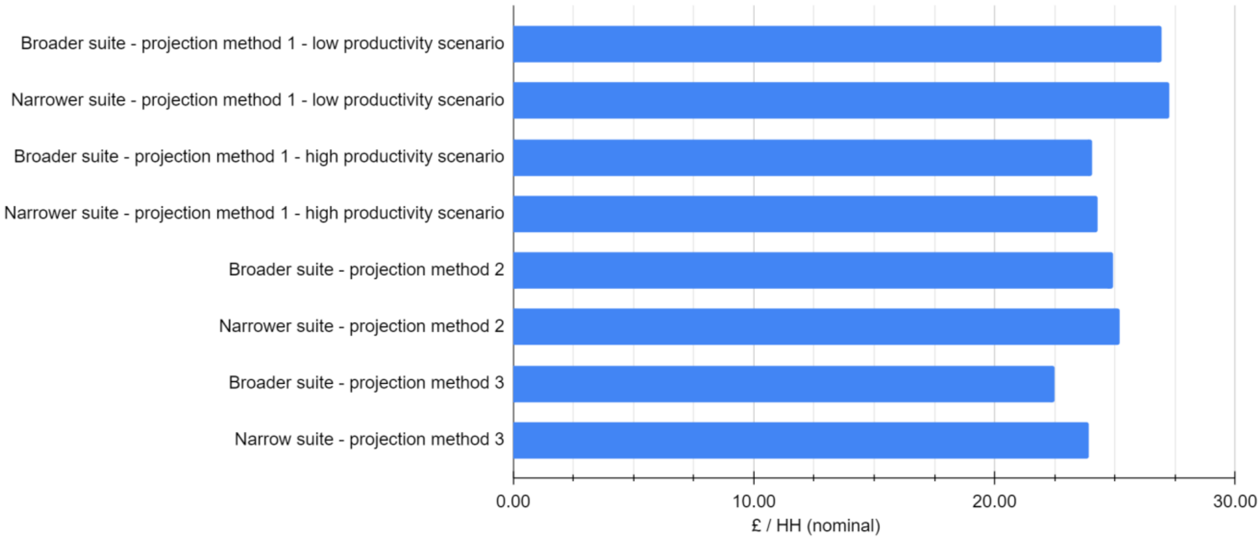


Figure 7 Projected efficient cost benchmarks for Wessex Water (nominal)



The projections for Bristol Water and Wessex Water are not directly comparable, not least because those for Bristol Water were based on some forecasts for customer numbers and meter penetration over the 2025-30 period whereas these forecasts were not available from Wessex Water and we applied forecasts for 2024/25 from Ofwat's PR19 FD for these variables over the 2025-30 period. As such, for example, the projections assume that Wessex Water's meter penetration rate remains at 0.67, the value forecast for 2024/25 at PR19. In contrast, the figures Bristol Water provided to us forecast significant increases to meter penetration over the period and this had a significant upward effect on its projected cost benchmarks.

Discussion of projected benchmarks and implications for PR24

At this stage in the PR24 process, we consider that the projection methods, and the approach and adjustments we have used to apply them, are more important than the precise values we calculated for the cost benchmarks over the 2025-30 period. This is due to several factors, such as:

- All the projections rely on OBR forecasts of inflation which are subject to particular uncertainty in the current abnormal inflationary environment. The OBR forecasts could change significantly over time. Furthermore, when companies prepare their business plans, they may use different inflation forecasts and there is scope for inconsistency with assumptions used for the projections presented in this report.
- As the PR24 process progresses, companies will have updated forecasts for some of the cost driver variables used to make the projections. As highlighted above, in relation to Bristol Water's meter penetration forecasts, these forecasts can have a significant influence on projected costs.
- The econometric models used data up to 2021/22 and data on water companies' outturn costs for 2022/23 will be available before companies finalise their business plans.

There may be benefit from an update in light of emerging data in these areas. Nonetheless, we provide some comments on the projections as they stand.

We can see from the charts set out above that the choice of model suite has some impact on the projected cost benchmarks, but this is generally quite small. The choice of projection method (and, for method 1, the assumed productivity scenario) has a somewhat greater impact. On this, the following points seem particularly relevant to highlight:

- Our view is that the productivity growth for an efficient water retailer over the period 2021/22 to 2029/30 is likely to lie somewhere between the low and high scenarios we use for projection method 1. The method 1 high scenario reflects a very approximate estimate of productivity growth of a notional efficient water retailer over our historical data period. At close to 3% per year, this is an unusually high rate of improvement and seems likely to reflect a phase of relatively high productivity improvements across the industry which resulted from a combination of some degree of historical inefficiency in the industry and the reforms to the regulatory approach to retail activities from PR14 onwards (e.g. separate controls targeted on residential retail activities, emphasis on benchmarking to set allowances, companies bearing more financial exposure to their bad debt costs). We have observed greater cost reductions in bad debt costs than in other retail costs which is consistent with the view about the impact of the PR14 reforms: companies had limited financial incentives on bad debt prior to PR14. We would not expect a productivity trend of this nature to be sustained over the long term. At the same time, it does not seem appropriate at present to base projections entirely on the 1% productivity trend that is often assumed in price control determinations and attributed to sources such as EU KLEMS. That value is well below what seems to have been achieved historically by water retailers and there does not yet seem to be evidence of a major slowdown.
- A similar point applies to the figures from projection methods 2 and 3 as for projection method 1 (high scenario). The real-term reductions in retail unit costs are likely to reflect a phase of relatively high productivity improvements following the PR14 reforms and we would not expect the cost trends (relative to CPIH) experienced historically to be sustainable over the longer term.
- Projection method 2 has some benefits over method 1 (high scenario) in that it does involve an attempt to decompose observed cost trends between productivity and input price effects, something which is quite an approximate and assumption-driven exercise. However, method 2 relies on an implicit assumption that the relationship between CPIH and water retailers' unit costs will be the same in the future than the past, whereas method 1 allows for greater flexibility on this. There may be some benefits from method 1 in using OBR forecasts of wage growth as well as CPI growth.
- Both projection method 1 (high scenario) and projection method 2 have the benefit that they use evidence on modelled cost trends for a notional efficient upper quartile retailer over the historical data period. In contrast, projection method 3 involves extrapolation of time trend coefficients from the econometric models which are not focused on unit cost changes for a notional efficient company. At present, we suggest that less weight is placed on results from projection method 3 as it does not seem as relevant to what we are interested in, and we have not identified significant benefits of it relative to methods 1 and 2 (perhaps other than simplicity in its application, which is less relevant for this report given our decision to apply all three methods).

In light of these considerations, and if a combined projection were to be needed, we can see a case for something along the lines of a weighted average which gives 50% weight to figures from projection method 1 (low scenario) and 25% weight to figures from projection method 1 (high scenario) and 25% weight to figures from projection method 2. This would weight equally productivity/unit cost evidence from inside and outside of water retailers and allow for evidence from an approach involving productivity and input price assumptions and evidence from an approach involving unit cost trend assumptions. But there is substantial uncertainty and imprecision in seeking to produce a point estimate for each year of the 2025-30 period.

More generally, the past is not necessarily a reliable guide to the future and there may be other factors to take into consideration if the estimates above are drawn on to make projections for the 2025-30 price control period.

Appendix 1: Review of Ofwat's PR19 approach to residential retail cost assessment

Introduction

We provide in this appendix a summary and review of a number of aspects of Ofwat's PR19 approach to residential retail cost assessment. We take the following topics in turn:

- Ofwat's overall approach to residential retail cost assessment.
- Comparison of outturn costs with PR19 allowances.
- Ofwat's approach to efficiency challenges and input prices for wholesale controls.
- Ofwat's use of business plan forecasts to inform its notional efficient costs adjustment for wholesale expenditure at PR19.
- The CMA's adjustment for notional efficient costs for wholesale expenditure.
- The starting point for productivity and input price adjustments under Ofwat's approach to wholesale controls.
- Ofwat's approach to the efficiency and input price adjustments for retail at PR19.
- The lack of adjustments for ongoing productivity and input price changes.
- The absence of analysis of evidence on trends in water industry retail costs.

Ofwat's PR24 final methodology does not say much about the issues covered in this appendix. We did not identify anything in the final methodology that indicates a major change of approach for residential retail cost assessment at the broad methodological level.

Ofwat's overall approach to residential retail cost assessment

We first summarise Ofwat's high-level approach to residential retail cost assessment, by which we mean the process and analysis used to set price control allowances for companies' residential retail costs (excluding issues relating to the costs of finance)[9].

Ofwat's approach to residential retail cost assessment at PR19 can – at a high level – be seen as involving the following broad approach:

- **Econometric benchmarking of historical retail costs in the water industry**. Ofwat estimated a suite of econometric benchmarking models, applied to historical data from water companies over a six-year period. These models included explanatory variables to try to take account of a range of cost drivers of retail costs.
- **Calculation of modelled costs**. For each company, in each year over the 2020-25 price control period, Ofwat calculated a measure of modelled unit costs (or modelled costs per

[9] In particular the determination of the retail margin allowed for residential retail finance costs in Ofwat's calculation of retail price controls and the calculations to separate the wholesale WACC from the appointee WACC.

household) as the predicted values from its suite of econometric models based on a set of forecasts for the 2020-25 period for each of the explanatory variables in the econometric models. It then multiplied the modelled unit costs by the corresponding company forecasts of customer numbers over the 2020-25 period to get figures for modelled costs in £m rather than on £ per customer basis.

- **Application of efficiency adjustment**. Ofwat assumed that an efficient water company would have a level of retail cost efficiency equal to the upper quartile level of retail efficiency across the industry. It applied an industry-wide efficiency adjustment of around -15% to the modelled costs calculated for each company. This was based in part on comparisons between its modelled costs and companies' business plans - with a number of companies planning to spend considerably less than the modelled costs implied - as well as on comparisons between companies' actual spend over the last five years of data and its modelled costs for those five years.
- **Cost adjustments**. Ofwat considered cost adjustment claims submitted by water companies in respect of residential retail, as a means to account for factors affecting the efficient level of costs for a company which are not adequately captured in the econometric benchmarking or other aspects of Ofwat's cost assessment. Ofwat made adjustments for only one of these claims at PR19, which concerned differences between England and Wales in the data on deprivation used in the econometric modelling.

While this process produced an allowance in £m for each company for each year of the 2020-25 price control period, the key metric of the price control for residential retail is the allowed cost per customer (connected household). The allowance in £m will adjust according to outturn customer numbers. Ofwat describes its retail control as an average revenue control, although technically it might be seen as a total revenue control with an adjustment mechanism based on differences between the forecast and actual number of customers supplied.

For PR19, Ofwat set the residential retail control in nominal terms, with no provision for CPIH-indexation or other inflation-related uncertainty mechanism.

Ofwat's PR24 final methodology does not indicate any major changes from the broad approach outlined above.

The approach to residential retail cost assessment applied at PR19, and expected for PR24, places a very large emphasis on high-level econometric cost benchmarking models to set allowances for retail costs. There is greater emphasis on econometric benchmarking than for wholesale cost assessment, for which some of the enhancement costs are assessed in a more evaluative way based on business plan justifications, and for which there are some categories of costs which are treated as unmodelled and excluded from cross-company benchmarking. For residential retail, the benchmarking is applied to all retail costs with the exception of pension deficit repair costs, third party costs and the costs of finance.

Furthermore, Ofwat's approach drew on benchmarking of forecast costs in business plans to set the efficiency adjustment which it did not do for wholesale (at least not as explicitly) which can be seen as a form of benchmarking to companies' projections of how their costs will evolve over time.

While water companies may have concerns about the emphasis on benchmarking, Ofwat's use of benchmarking seems reasonable, especially given the positive incentive properties of a benchmarking approach to cost assessment and the view that water companies' retail costs should be relatively amenable to benchmarking.

In terms of improving the approach to retail cost assessment at PR24, we see opportunities for improvements to be made in the following key areas:

- **Econometric benchmarking**. Ofwat's suite of econometric models, including the specification of individual models and the balance of modelling approaches across the suite of models used.
- **Indexation**. Whether and how the residential retail price control is adjusted according to an economy-wide inflation index (e.g. CPIH) or according to changes in some measure of outturn input prices or costs relative to assumptions made when setting the price control.
- **Moving from modelled costs to projected cost benchmarks**. The way that the notional efficiency company efficiency adjustment is determined, and the consideration of factors such as ongoing productivity improvements, input price changes over time and the cost implications of expected improvements in service quality when moving from econometric models estimated using historical data to setting allowances for the forthcoming price control period.

The remainder of this appendix is focused on the third area above.

There may also be scope for improvements around the review and treatment of cost adjustment claims, but this is not a priority for the project and is not covered in this report. We have not sought to consider potential cost adjustment claims that Wessex Water and/or Bristol Water might make.

Comparison of outturn costs with PR19 allowances

Before considering various aspects of Ofwat's PR19 cost assessment approach in more detail, we provide some comparison of outturn costs for 2020/21 and 2021/22 with the cost allowances set by Ofwat at PR19. Figure 8 presents the over- or under-spend of each company compared to PR19 residential retail allowances, taking 2020/21 and 2021/22 together. The figures are on a nominal basis.

As shown, across the industry, companies' outturn costs in each of the first two years of AMP7 have tended to be substantially above the cost allowances set at PR19 for those two years. Across the industry, companies' outturn costs across those two years exceeded allowances by 19%, though, as shown in the chart, such an average figure masks considerable variation across companies. Southern Water, Welsh Water and Sutton and East Surrey were amongst those whose costs, in percentage terms, most exceeded their allowances; by more than 40% for all three companies. On the other hand, the out-turn costs for Anglian Water, South West Water and Hafren Dyfrdwy were below those companies' PR19 allowances.

Figure 9 below provides some additional detail, giving a breakdown of unit costs by cost categories. For the purpose of this chart, the category labelled "Other operating expenditure" reflects the aggregate of the following different APR cost lines: local rates and cumulo rates, net recharges and their operating expenditure.

Figure 8 Outturn costs vs cost allowances (2020/21 – 2021/22)

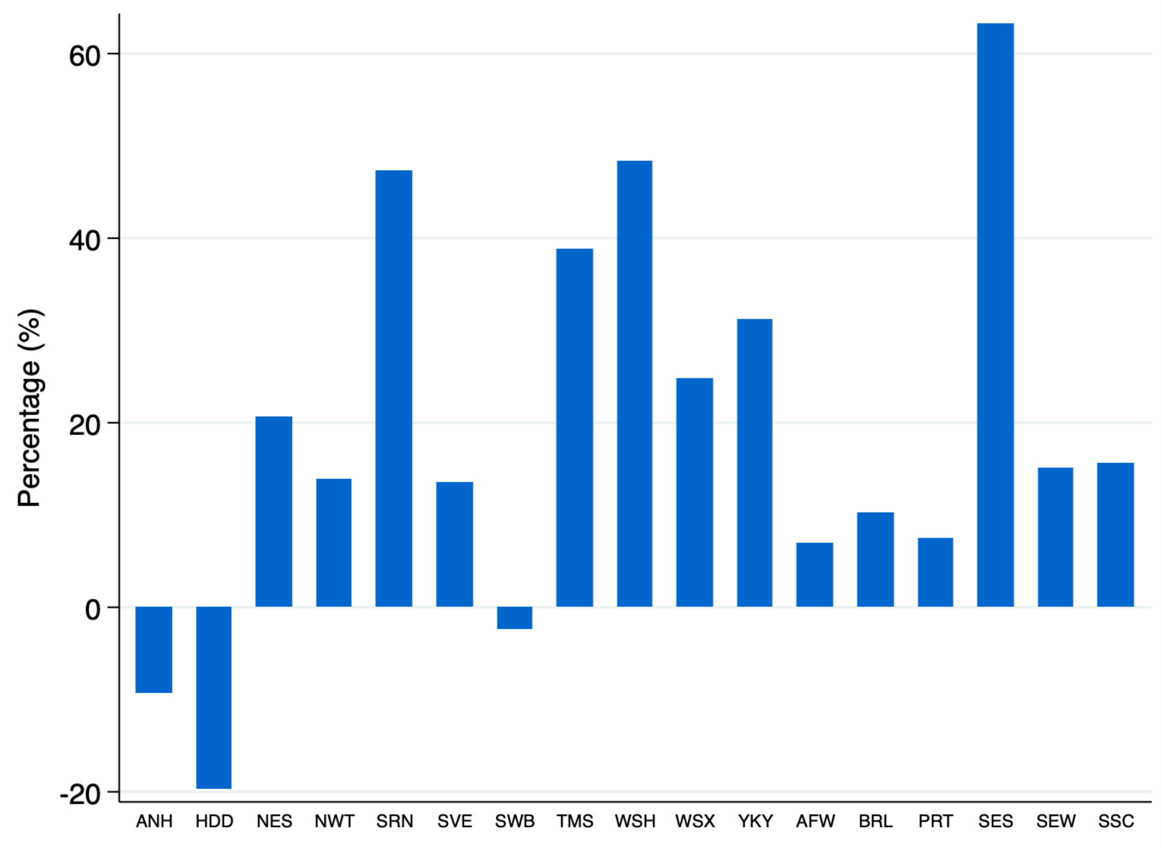


Figure 9 Break-down of industry-average retail cost per connected household: (2013/14 – 2021/22)

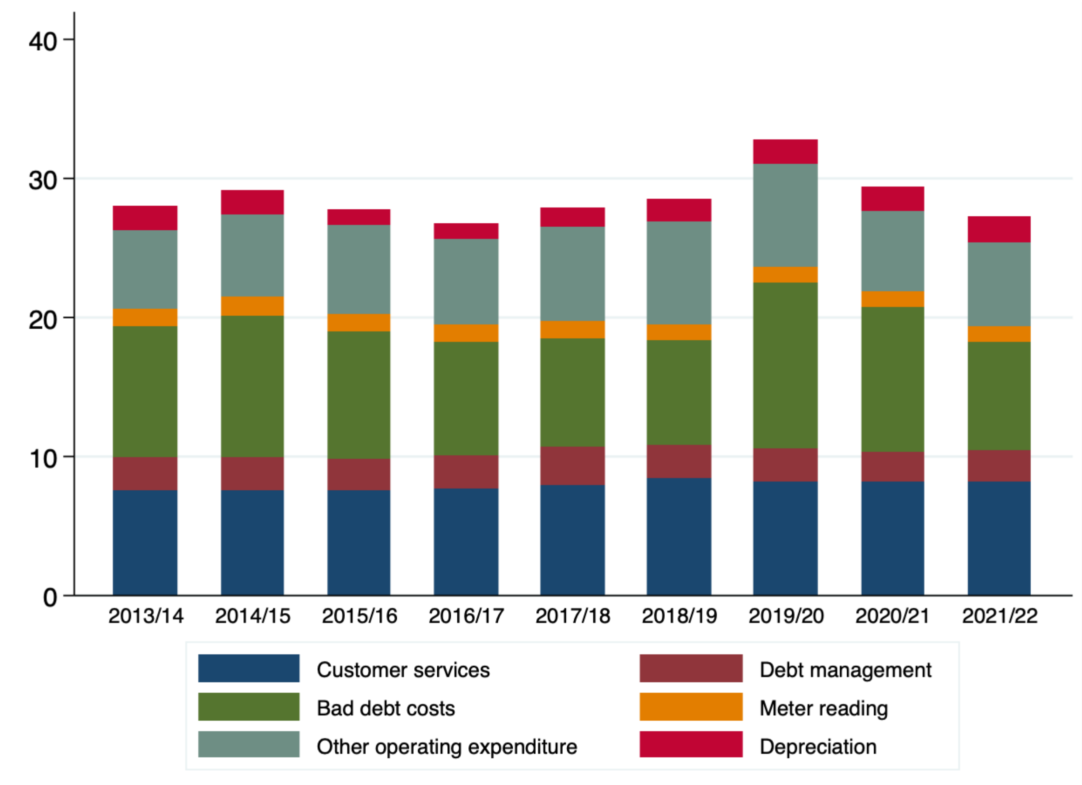


Figure 9 illustrates that the significant increase in unit costs from 2018/19 to 2020/21 was due to the higher costs associated with the provisions for bad debts, something which companies' commentaries in their APRs explain by reference to concerns about the Covid-19 pandemic.[10] Further to that, the chart also illustrates that, in broad terms, and making an allowance for the increase bad debt costs potentially associated with Covid-19, the relative share of costs accounted for by the different cost categories have remained fairly stable over time.

Ofwat's approach to efficiency challenges and input prices for wholesale controls

Part of Ofwat's approach to residential retail cost assessment at PR19 was rather novel and it is useful to compare this to the approach taken to wholesale costs, which was more conventional and consistent with wider regulatory precedent. On this basis, we first consider relevant aspects of the approach Ofwat applied to wholesale costs at PR19.

For the purposes of projecting cost benchmarks over the 2020-25 period, Ofwat's approach to wholesale price controls at PR19 involved three types of adjustment being applied to its calculation of modelled costs for each company over that period.[11] We summarise these in the table below. In this table we refer to both the terminology used by Ofwat as well as the terminology we used in section 2 of this report.

Table 9 Adjustments for efficiency and input prices applied at PR19 for wholesale costs

Adjustment type	Brief explanation
Adjustment for differences between notional company efficiency and industry-average efficiency Ofwat refers to this as an adjustment for catch-up efficiency or as a catch-up (efficiency) challenge	• This is intended to adjust the cost benchmarks for each company derived from the econometric modelling so that these are consistent with the level of efficiency of what Ofwat considers to be a notional efficient company over some historical period (e.g. over the historical period used for the benchmarking or a shorter sub-sample of this such as the last five years). • The adjustment is calculated by making comparisons between (a) each company's actual costs aggregated over the selected historical period and (b) the cost benchmarks/predictions (or modelled costs) for each company derived from the econometric modelling, aggregated over this period. • The company with the lowest ratio of actual costs to modelled costs is deemed to be the most efficient company and this ratio is used to rank the efficiency other companies accordingly. • The notional efficient company might be assumed to be a company operating at the estimated upper quartile level of cost efficiency (based on these ratios) or some other percentile. • The value of adjustment is the same for all companies and reflects the ratio between actual costs and modelled costs for the notional efficient company (e.g. the upper quartile company) – for example, if the notional efficient company's costs were 97 and its modelled costs 100, the

[10] These figures use companies' reported bad debt costs before application of the smoothing adjustments that Ofwat subsequently asked companies to provide.

[11] We follow Ofwat's convention in using the term modelled costs to refer to the £m cost predictions for a given water company that are based on the predicted values from the econometric models. At PR19 the modelled costs for the 2020-25 period were calculated using forecast explanatory variables (and forecast customer numbers) for that period, and represent cost benchmarks derived directly from the econometric models, before the application of any adjustments for upper quartile efficiency, ongoing efficiency and input prices.

Adjustment type	Brief explanation
	adjustment would be a 3% deduction from modelled costs applied to all companies. • For wholesale water at PR19 Ofwat used the company it deemed to be fourth most efficient (out of 17) as the notional efficient company for the purposes of its adjustment for notional efficient costs, and for wholesale wastewater it used the third most efficient (out of ten). This led to a 4.6% reduction applied to modelled wholesale water costs and an 8.7% reduction applied to wholesale wastewater costs.
Adjustment for ongoing productivity improvements of a notional efficient company Ofwat referred to this as its frontier shift (efficiency) challenge or frontier shift (efficiency) adjustment	• While the adjustment summarised in the row above might be seen to reflect historical levels of efficiency for a notional efficiency company, this adjustment is intended to allow for the improvements in productivity that a notional efficient company would be expected to make in the period to the end of the forthcoming price control period (e.g. to 2024/25 in the case of PR19) relative to its historical levels of productivity/efficiency. • The same adjustment was made for all companies in the industry. • This adjustment is not calculated mechanistically and is a regulatory judgement informed by a range of evidence and arguments. • For wholesale water and wastewater at PR19 Ofwat set this figure as 1.1% per year applied from 2019/20 onwards. This figure was considered during the PR19 appeals to the CMA and the CMA used a figure of 1% per year instead of 1.1%.
Adjustment for changes to the input prices faced by a notional efficient company Ofwat referred to this as an input price pressure adjustment; an adjustment/allowance for real price effects; and an adjustment for (real) input price inflation.	• This adjustment is intended to allow for the changes in input prices that a notional efficient company would be expected to face in the period to the end of the forthcoming price control period (e.g. to 2024/25 in the case of PR19) relative to the levels of input prices • For wholesale at PR19, given the indexation of wholesale price controls, this adjustment is for changes in input prices relative to the price control inflation index or in "real" terms, rather than for nominal input price changes. • The same adjustment was made for all companies in the industry. • This adjustment is not calculated mechanistically and is a regulatory judgement informed by a range of evidence and arguments. • For PR19 at the wholesale level, the adjustment for input price changes was around 0.4% to 0.5% per year (relative to CPIH).

While terminology may differ, we consider that the broad approach to the three elements above fits with established regulatory precedent, including the CMA's approach for the PR19 appeals and Ofgem's approach for energy network price controls. We consider the first of adjustment in a little more detail below as it is particularly relevant for comparisons with the approach to retail costs at PR19.

Ofwat's use of business plan forecasts to inform its notional efficient costs adjustment for wholesale expenditure at PR19

While Ofwat's adjustment for notional efficient costs based on historical cost levels is ostensibly an adjustment based on evidence on historical performance, in practice Ofwat took account of forward-looking evidence too – specifically water company business plans – in deciding what company to use as the notional efficient company for the purposes of calculating the adjustments for wholesale water and wholesale wastewater at PR19.

Ofwat's decision on the choice of which company to use as the benchmark for a notional efficient company involved the use of the upper quartile company for draft determinations but for final determinations Ofwat set a more demanding benchmark. In making that revision, Ofwat seems to have been influenced by two main considerations:

- A view that the adjustments calculated at final determinations using updated data and the upper quartile company as the notional efficient company were not large enough or not challenging enough in their own right. Ofwat reported that the adjustments using an upper quartile benchmark for final determinations would imply a 3.9% reduction to modelled costs for wholesale water and a 1.2% reduction to modelled costs for wholesale wastewater. This compared to reductions of 6.5% and 10.4% respectively at the PR14 final determinations and reductions of 4.8% and 3.7% respectively at the PR19 initial assessment of business plans stage.
- Comparisons of the modelled costs that would apply under different assumptions for what company to take as the notional efficient company (i.e. upper quartile or something more demanding) against companies' updated business plan forecasts for modelled costs over the 2020-25 period.[12]

Ofwat explained this as follows:[13]

> *"Following changes to our data and modelling approach since draft determinations (e.g. the removal of non-section 185 diversions costs and the inclusion of the 2018-19 data), the stringency of the historical upper quartile as a catch up efficiency challenge has reduced. Out of 17 water companies, 12 companies now forecast modelled base costs for the period 2020-25 that are lower (i.e. more efficient) than the projected efficient costs under the historical upper quartile. This compares to only six out of 17 at the slow track draft determinations. The historical upper quartile does not appear to deliver a strong challenge for the sector at final determinations. We acknowledge that part of the reason for the reduced challenge is companies reducing their requested costs in August 2019 representations to draft determinations. This may reflect that companies have improved the understanding of their costs through the price review process. …*
>
> *Setting the challenge at the fourth placed company in water and third placed company in wastewater still leaves eight out of 17 companies whose business plans' modelled base costs are more efficient than our efficient benchmark. Moreover, it leaves sufficient headroom relative to the most efficient companies in the sector. We consider that our catch-up challenge remains on the conservative side and could have been stretched even further. However, taking into account the overall stretch of our determinations, we consider this choice to be appropriate and in the interest of customers."*

As Ofwat acknowledges, part of the reason for why the adjustment based on an upper quartile benchmark reduces between initial assessment of business plans phase and the final determinations phase is that companies reduced their cost forecasts. In this context it would not be

---

[12] We have not sought to check the veracity of Ofwat's comparisons in this respect (e.g. whether on a like-for-like basis).

[13] Ofwat (2022) *PR19 final determinations: Securing cost efficiency technical appendix*, pages 32-34.

reasonable, we think, for Ofwat to target the same scale of adjustment for notional company costs in percentage terms at final determinations as was calculated at the initial assessment of plans stage.

The key point for the purposes of this section is that, even for wholesale cost assessment at PR19, Ofwat seems to have been influenced by comparisons of water companies' business plan forecasts. But this was much less explicit than for the retail cost assessment.

In its PR24 final methodology, Ofwat said that it intends "*to use a combination of historical and (where appropriate) forecast efficiency evidence to set the catch-up efficiency challenge at PR24"*, in the context of both wholesale and retail controls".[14]

The CMA's adjustment for notional efficient costs for wholesale expenditure

The main focus of this section is on Ofwat's approach at PR19 and we have not reviewed the CMA's determination, not least because the CMA did not consider the retail controls. Nonetheless, we briefly highlight below the CMA's reasoning on the choice of the notional efficient company for what it described as the catch-up efficiency challenge in the context of its determinations of wholesale controls.

For wholesale water and wholesale wastewater, the CMA used an upper quartile benchmark for the notional efficient company, without any reference to business plan forecasts to support this. This resulted in a reduction of 1.4% being applied to modelled costs for wholesale water and a 2.2% reduction being applied to modelled costs for wholesale wastewater. The CMA had been able to use an extra year's data in its econometric modelling, and calculated the adjustment based on analysis of relative efficiency of companies over the most recent five-year period.

The CMA summarised the reasoning for its decision as follows:[15]

> *"Our cost models estimate how much it would cost the averagely efficient water company to cover base operations. However, we want to set cost allowances for a water company that is more than merely averagely efficient, and so we apply a 'catch up' efficiency challenge. Our decision is to use the company at the upper quartile as the benchmark and reduce the Disputing Companies' allowances accordingly. We consider this sets a challenging benchmark whilst acknowledging the limitations of our econometric modelling (and the consequent risk that the company will have insufficient allowed revenue to ensure a base level of service). Our benchmark is set at a similar, although slightly less demanding, level to Ofwat's."*

The CMA also explicitly listed a number of factors that it had placed "little or no weight on" but which had been raised as part of the appeals process. Its comments on three of them seem particularly relevant. The CMA explained as follows:[16]

---

[14] Ofwat (2022) *Creating tomorrow, together: Our final methodology for PR24: Appendix 9 – Setting expenditure allowances*, page 36.

[15] CMA (2021) *Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations: final report*; page 19.

[16] CMA (2021) *Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations: final report*; pages 231-232.

"[we placed little or no weight on:]

… the fact that one regulator, Ofgem, had chosen an efficiency challenge 'tougher' than upper quartile. These regulators are regulating different sectors with different companies, so there is limited read across to our decision.

… evidence that the absolute level of the efficiency challenge had fallen, particularly for wastewater. We found that it was more appropriate to set the efficiency challenge based on our assessment of the quality of the econometric modelling, rather than to seek specific outcomes.

… a comparison of the companies' business plans with the modelled allowances. We found it was more appropriate to set the efficiency challenge based on our assessment of the quality of the econometric modelling, rather than to seek specific outcomes."

The CMA, therefore, seems to have explicitly decided against following Ofwat's approach of using comparisons with business plan forecasts to inform the notional efficient company adjustment, at least for wholesale costs which were the focus of its determination.

The starting point for productivity and input price adjustments under Ofwat's approach to wholesale controls

When applying adjustments for ongoing productivity improvements of a notional efficient company and input price changes, there is a question of what year these adjustments should be applied from. For example, if the net effect of these adjustments is taken to be a 0.5% decrease in (CPIH-adjusted) costs per year, then from what year should the 0.5% decrease start to apply?

For wholesale cost assessment, Ofwat applied the adjustments from 2019/20 and explained this as follows: [17]

"We will apply the frontier shift and real price effects to costs from 2019-20 rather than 2020-21. Base cost inputs, and cost forecasts, used in our cost models only take into account data and therefore on-going efficiency improvements and real price effects up until 2018-19. We therefore consider it appropriate to add in frontier shift and real price effects for the additional year."

To clarify, Ofwat's approach involved applying an annual adjustment of around -0.5% to modelled costs in order to calculate costs for 2019/20 and each subsequent year up to 2024/25.

Ofwat felt the need to explain why it was making the first adjustment for the 2019/20 financial year in preference to making the first adjustment for 2020/21 and its explanation for this specific choice seems reasonable. However our view is that, on closer inspection, it is entirely logical to apply the adjustment from 2019/20 and it would probably make more sense if it were applied from an earlier point in time.

The implicit assumption in the approach taken by Ofwat is that the cost benchmarks derived from its econometric modelling are representative of the level of input prices and productivity prevailing in

[17] Ofwat (2022) *PR19 final determinations: Securing cost efficiency technical appendix*, page 122.

the last year of its dataset: 2018/19, and that adjustments are then needed to move from costs on a 2018/19 basis to costs on a 2019/20 basis and so forth for each consecutive year.

This implicit assumption does not stand up to scrutiny in a context where:

- The sample period for its econometric modelling spanned six years and where all its model specifications involved only a constant term and no time trend, time dummies or other dynamic components. For this type of model, we do not see how the predicted values can reflect the level of input prices and productivity prevailing in 2018/19. Similar weight in the model estimation process will be given to all years from 2013/14 to 2018/19 and the models impose the assumption of no changes in CPIH-adjusted costs over time (other than changes explained by changes in cost driver explanatory variables).
- Ofwat's notional company efficiency adjustments are calculated using historical data over the last five years of the dataset.

Given the design of Ofwat's econometric models, its approach to adjustments for notional efficient costs, and the significant changes in retail costs observed across the industry over the historical data period, we consider it is highly unlikely that the benchmarks from Ofwat's econometric modelling would be representative of the input prices and productivity prevailing in 2018/19.

In its determination on the PR19 appeals, in the context of base cost assessment for the wholesale price controls, the CMA said that "*Our cost models estimate how much it would cost the averagely efficient water company to cover base operations*". We think that the CMA's statement is reasonable, but it leaves open the question of whether its models estimated the costs of the averagely efficient water company in the last year of its data (2019/20) or the averagely efficient water company over the full sample period used for the modelling. The CMA's logic would imply the former rather than the latter. But, like Ofwat, the CMA applied its adjustments for ongoing productivity (frontier shift) from the first year after the end of its econometric sample period (so the CMA applied it to 2020/21 and subsequent years as its econometric models used data including 2019/20).

We suggest that a more valid position to take – at least as a starting point in the absence of evidence to the contrary – is that:

- If the econometric models have constant terms and no time trend or time dummy variables, the econometric model estimation process will result in estimated cost benchmarks for each company that reflect the average level of productivity over the sample period of six years and the average level of input prices over that period.
- If a notional efficient company adjustment (e.g. upper quartile) is calculated on a recent subperiod (e.g. the last five years rather than the last nine years) the application of this will tend to update the productivity and input prices implicit in the cost benchmarks or modelled costs to the averages over that subperiod.
- In the absence of evidence on how productivity and input prices across the industry have evolved during the sample period or subperiod, the level of productivity and input prices that

applied on average over that period/subperiod might reasonably be approximated as the level applying in the mid-year of that period/subperiod.

Turning to practicalities, on this basis we see an argument that adjustments for frontier shift and real price effects at PR19 would have made more sense if they had been applied from 2017/18 onwards rather than 2019/20 onwards (i.e. compounded over a somewhat longer period when being used to make adjustments to modelled costs). However, that rests on the econometric models having a constant term only, and we would generally expect the econometric models for PR24 to warrant either a time trend, time dummy variables or some combination of the two.

While this may seem a minor point, we wonder whether it helps to explain Ofwat's reluctance to use the conventional approach – as used for wholesale controls – when it came to the retail controls. Water companies' residential retail costs have generally been falling over time on a CPIH-adjusted basis. In this context, the conventional approach will tend to over-estimate the level of companies' efficient costs over the forthcoming price control period.

While Ofwat did not articulate any of the technical issues above, it may have been mindful that a set of benchmarking models and adjustments based on the last five or six years of historical data is unlikely to provide an up-to-date view of retail costs which could be taken as a starting point for the 2020-25 period. This, in turn, may have encouraged it to place more weight on company forecasts.

Ofwat's approach to the efficiency and input price adjustments for retail at PR19

We now turn to Ofwat's approach to the efficiency and input price adjustments for retail costs at PR19. We reproduce in detail Ofwat's PR19 description of its approach to the efficiency challenge – which also relates to ongoing productivity improvements and input price changes – as it was novel and is important for the current project:[18]

> *"We calculate the upper quartile efficient level of cost performance, which corresponds to the level of costs that the top 25% of companies achieve. We apply the average of the historical and forward-looking upper quartile efficiency challenges to companies' modelled costs to set allowances for 2020-25.*
>
> *"Our approach to applying the efficiency challenge for residential retail differs to that in wholesale. In retail, we apply the average of the historical catch-up and the forward-looking upper quartile efficiency challenges to companies' modelled costs over 2020-25 to calculate PR19 allowances. We consider that using business plans to inform the efficiency challenge is appropriate, particularly for retail services. The retail control has started as recently as 2015 and retail services can transform more quickly than wholesale services (eg due to lack of long-lived infrastructure assets). The fact that the majority of companies submitted forecasts that are significantly more efficient than historical expenditure is evidence of the pace at which this service is transforming. It is important that customers share the benefits. We consider that the upper quartile benchmark, which is determined by the fifth most efficient business plan, provides a credible challenge for the sector.*

---

[18] Ofwat (2019) *PR19 final determinations: Securing cost efficiency technical appendix*, page 128.

"We do not apply a further frontier shift challenge or input price pressure adjustment in residential retail. Efficient business plans may have accounted for these in their cost forecasts, which feed into our forward-looking efficiency challenge, and thereby are reflected in our efficient allowances."

Ofwat also elaborated on how it saw its approach as setting allowances in a way that combines evidence on historical performance with evidence from business plans:[19]

"We maintain our view that it is appropriate to calculate the efficiency challenge as the average of the historical catch-up and the forward-looking efficiency challenges. We consider that setting a catch up challenge which puts 50% weight on historical cost performance and 50% on business plan performance strikes an appropriate balance of evidence in the more dynamic retail sector."

Ofwat also highlighted its view that its retail cost challenge was reasonable in light of some company forecasts and its analysis of trends in costs (though Ofwat did present such analysis in its document):[20]

"Furthermore, our analysis reveals that our cost challenge at PR19 is planned to be achieved or exceeded by seven companies. Our analysis shows that some companies have already achieved this level of performance in recent years, while for others, the expenditure path shows a clear trend towards our allowances. However, for a small number of companies our cost challenge in residential retail remains significant. This is the case for SES Water, Wessex Water, Dŵr Cymru, Northumbrian Water and Thames Water. We expect these companies to catch up to efficient companies in the sector. SES Water, Wessex Water and Northumbrian Water were among the least ambitious in terms of the cost reductions they proposed in their PR19 business plans relative to actual historical spend. SES Water, Dŵr Cymru and Thames Water are underperforming on their PR14 retail allowances in the current regulatory period (2015-2020) more than other companies, despite the glide path to catch up with the rest of the industry we allowed at PR14. We consider that these companies had sufficient time to catch up with efficient retailers in the sector and therefore maintain our view that our allowances provide an appropriate basis for setting cost allowances for the industry."

Ofwat's use of company forecasts had been signalled in Ofwat's PR19 final methodology. It had said the following about retail cost assessment: "*Our efficient cost baselines will be informed by historical and forward looking performance in the sector, as well as by relevant information from outside the sector*".[21]

Drawing on Ofwat's explanation above, and our review of the Excel file produced by Ofwat to show its calculations, we can look in more detail at Ofwat's approach. The approach involved making adjustments to the modelled costs in two different ways such that it calculated two different cost

[19] Ofwat (2019) *PR19 final determinations: Securing cost efficiency technical appendix*, page 131.

[20] Ofwat (2019) *PR19 final determinations: Securing cost efficiency technical appendix*, page 131.

[21] Ofwat (2017) *Delivering Water 2020: Our final methodology for the 2019 price review: Appendix 11: Securing cost efficiency*, page 18.

benchmarks, which it then gave equal weight (50% each) in calculating allowances for the 2020-25 period. Ofwat's approach produced two sets of cost allowances:

- **Component (A): Adjustment for notional efficient costs based on comparisons of modelled costs with business plan forecasts**. For this component, Ofwat adapted the type of approach used for wholesale to calculate an adjustment for the costs of a notional efficient company so that it was based on business plan cost forecasts rather than historical costs. For each company Ofwat calculated the ratio of the company's business plan forecasts for residential retail costs expressed in nominal terms (aggregated over the five-year price control period) and Ofwat's modelled costs for that company over the same period (taking account of forecasts explanatory variables). The company with the smallest ratio of forecast costs to modelled costs was deemed the most efficient and other companies ranked accordingly. Ofwat selected the upper quartile company as the most efficient company. On this basis, Ofwat calculated a reduction of 20.6% to modelled costs for each company for the 2020-25 period, reflecting the upper quartile company's business plan forecast costs being 20.6% lower in nominal terms than Ofwat's modelled costs for that company, which were expressed in 2017/18 prices.
- **Component (B): Adjustment for notional efficient costs based on historical cost levels**. This component involves the familiar type of adjustment for notional efficient costs, as described above for wholesale costs at PR19, based on comparisons of companies' actual costs of the last five years of data relative to modelled costs for the same period. On this basis, Ofwat calculated a reduction of 10.6% to modelled costs for each company for the 2020-25 period.

Taking these two together, and weighting them equally, Ofwat applied an adjustment of 15.4% to modelled costs for the 2020-25 period. This is a large adjustment compared to what we saw for wholesale costs at PR19.

For neither of the two components summarised above did Ofwat apply any form of adjustment for (a) ongoing productivity improvements of a notional efficient company (frontier shift efficiency challenge) or (b) input price changes. This is in stark contrast to the approach Ofwat used at PR19 for wholesale controls, the approach taken by the CMA (e.g. for the PR19 appeals) and that taken by Ofgem for energy network price controls.

Ofwat's approach on this issue seems valid in respect of component A. Given that Ofwat used companies' business plan forecasts (in nominal terms) for the comparisons used to calculate the adjustment applied to modelled costs, it seems appropriate not to make any further adjustments for productivity improvements over time or for input prices, as these should be reflected in companies' forecasts. Ofwat was not necessarily bound to accept the assumptions on ongoing productivity improvements over time or for input prices that were implicitly or explicitly used by companies to produce their forecasts. But Ofwat could reasonably argue that the allowances arising from component A already take account of ongoing productivity improvements and anticipated nominal input price inflation.

Ofwat's approach on this issue seems to be mistaken in respect of component B. The adjustment for component B reflects outturn historical costs only, and does not take any account of ongoing productivity improvements, or input price changes, between the historical period used for Ofwat's modelling and notional efficient company adjustment and the 2020-25 period.

Ofwat's explanation does not make sense. As highlighted above, Ofwat explained the lack of a frontier shift challenge or input price pressure adjustment by saying that efficient business plans may have accounted for these. But in setting retail allowances, Ofwat only gave 50% weight to figures that reflected business plan forecasts. The remaining 50% weight, given to what we have called Component B, is based on historical modelling only; for this Ofwat's explanation for not considering frontier shift challenge or input price pressure is not applicable.

Our tentative view is that Ofwat made an error in this part of its PR19 cost assessment, overlooking in particular the implications of the lack of inflation indexation for the calculations of costs needed for its component B adjustment.

Interestingly, one of Ofwat's spreadsheets published as part of its final determinations includes assumptions and calculations that enable figures to be produced which do apply assumptions for industry-wide productivity improvements (frontier shift) and input price pressures. These seem to be the same assumptions for frontier shift and input price pressures: 1.1% per year and 0.4-0.5% per year for input price pressures which were used for wholesale. We do not think that Ofwat used these figures directly for the calculation of the final allowances, but they may have informed its thinking and judgement. However, the calculations made on this spreadsheet do not take account of the lack of indexation and apply an adjustment for real price effects rather than nominal price effects, which would act to underestimate the level of efficient costs needed for setting the retail price controls over the 2020-25 period. This lends additional credence to the hypothesis of a mistake in relation to the treatment nominal input price inflation for retail costs.

The absence of analysis of evidence on trends in water industry retail costs

One notable absence from Ofwat's retail cost assessment – as presented in its PR19 final determinations – is the analysis of changes over time in residential retail costs of the water companies it regulates.

At a simple level, Ofwat might have presented figures or a chart on the changes over time in the average level of retail costs per household in constant prices over the six-year period that it used for its econometric modelling.

Alternatively, Ofwat might have explored econometric models with time trends to see if these could provide usable evidence on the changes in unit costs over time, controlling for changes in cost drivers.

The efficiency adjustment applied to modelled costs of 15% seems demanding, and abnormally high in relation to wider regulatory precedent, especially given the lack of inflation indexation so that the implied adjustment in real terms (e.g. CPIH adjusted) would be even higher. It would have been informative to compare the assumptions on what is considered achievable by an efficient company over the 2020-25 period with what was achieved over the previous six years. This exercise is not straightforward, however, because industry-wide cost reductions in retail may reflect an element of catch-up and convergence and might not provide a direct guide to the ongoing productivity improvement of a notional efficient company over the historical period. This is something that we have sought to tackle in the analysis we carried out under our application of projection method 2 in section 3 of this report.

Appendix 2: Specification of projection methods

This appendix provides further information on the methods that might be used to take the results from the econometric benchmarking models of residential retail costs and to make projections of cost benchmarks over a forthcoming price control period.

In particular, we set out a number of methods – which we refer to as projection methods – that we consider to be internally coherent in the treatment of four key factors, relating to: movements over time in cost driver variables; differences between notional company efficiency and industry-average efficiency; the ongoing productivity improvements of a notional efficient company; and the changes to the input prices faced by a notional efficient company.

The methods we present include versions of those that Ofwat and the CMA used at PR19 for wholesale and retail controls. In addition to these, we identify some alternative methods that draw on different types of evidence and which could play a role at PR24, either in place of, or alongside, the types of method used for PR19.

These five methods are presented as potential options to consider. Which of them are reasonable to apply in a specific case will depend on the nature of the evidence available to apply them and on results from their application. Furthermore, while these methods are intended to be coherent in principle and at a methodological level, whether they are coherent in practice will depend on choices made about values for the various adjustments or assumptions and the evidence to support this.

When making projections of cost benchmarks – or setting allowances at a price control review - there may be merit in drawing on more than one of the methods (e.g. by using an average of implied cost allowance across multiple approaches), as any single method is likely to have benefits and drawbacks in the context in which it is applied.

We describe the five methods in more detail in Table 10. In the table, we set out for each method:

- The key steps or adjustments applied under the method to take account of the combined effects of: (a) movements over time in cost driver variables; (b) differences between notional company efficiency and industry-average efficiency; (c) the ongoing productivity improvements of a notional efficient company; and (d) the changes to the input prices faced by a notional efficient company).
- Which of these four factors are taken into account by each of these individual steps or adjustments. For each factor, a dark circle indicates that the effects of the factor can be taken into account by that step. For ongoing productivity and input price effects the dark circle concerns whether the effects between the historical data period and the end of the forecast period can be taken into account via the step (i.e. not whether the step may take account of effects of these factors during the historical data period used for the modelling).
- Brief comments on the sources of evidence that might be used to inform each individual step or and adjustment and on key areas of regulatory discretion or judgement.

Using an example to help clarify the interpretation of the table, we can see from the table that for the second method there are three key steps. Of these, the first takes account of factors relating to movements over time in cost driver variables, the second takes account of the notional efficient company, and the third step allows for the combined effects of ongoing productivity and input price changes.

The five broad methods described in the table represent what we see as a diverse and relevant set of methods to present in this report, but these are not exhaustive of all possible approaches, and are not necessarily the most appropriate approaches. There may further alternatives or variants on these to consider, such as:

- There is a potential variant on method 4 in which the use of company business plan forecasts is separated from the notional efficient company adjustment. For instance, this could involve first calculating a conventional upper quartile efficiency adjustment based on historical costs (as under steps 1B, 2B and 3B for other projection methods) and then calculating a further adjustment, for the net effects of productivity improvements and input price changes, as the *average* or *median* across companies of each company's ratio of (a) forecast costs in nominal terms to (b) modelled costs (without applying any efficiency adjustment to these modelled costs). To make projections for the 2025-30 period, both the historical upper quartile adjustment and the further adjustment for productivity and input price effects would be applied to modelled costs.
- There is a potential variant on method 3 in which the notional efficient company adjustment is calculated using comparisons of companies' business plan cost forecasts over the forthcoming price control period against modelled costs for that period (i.e. as for step 4B under method 4) rather than using comparisons of companies' actual costs over a historical period against modelled costs for that period. This type of approach might be seen as a hybrid of methods 3 and 4.
- There is a variant of method 5 in which companies' forecasts are included in the data feeding into the econometric benchmarking models, but where the company forecasts used are forecasts of retail costs for the forthcoming price control period under an artificial assumption of no ongoing productivity improvement over time and no changes in input prices (i.e. companies' explicit or implicit forecasts for ongoing productivity and input price changes are stripped out). Under this type of method there would then be a role for applying separate annual adjustment factors for ongoing productivity and input prices (e.g. the type used under approach 1) to the modelled costs over the 2020-25 period as these would not be reflected in forecasts. Ofgem used a version of this approach for gas distribution companies in its RIIO-GD2 determinations (and there was a dispute taken to the CMA concerning how Ofgem stripped out ongoing productivity from one of the company's forecasts). This type of approach might be seen as a hybrid of methods 1 and 5.

We have not sought to consider these variants further in this report. We wanted to keep the number of approaches presented as reasonably manageable and proportionate.

Table 10 Explanation of projection methods

Projection method and key steps/adjustments involved	Factors captured by step				Comments
	Cost driver variables	Notional efficient company	Ongoing productivity	Input price changes	
1. Application of separate productivity and input price assumptions					
Step 1A. Modelled cost over forthcoming price control period estimated using econometric benchmarking models.	●	○	○	○	Modelled cost for forthcoming price control period draws on forecasts for explanatory variables and unit cost denominator over that period. Otherwise a mechanistic calculation based on estimation results from the suite of econometric models. If models have time trends variables these should not be applied beyond the end of the historical data period.
Step 1B. Conventional notional efficient company adjustment based on comparisons of companies' actual costs over last X years against their modelled costs for that period.	○	●	○	○	Regulatory judgement on choice of notional efficient company (e.g. upper quartile or more demanding), which might be informed by range of considerations such as views on model quality and potentially comparisons of projected modelled cost post adjustment versus companies' forecasts. Regulatory judgement on what historical time period to use for comparisons of historical versus modelled costs (i.e. X in step 1B). A shorter X will allow for benchmarks to be more up-to-date though possible concerns that using a small number of years might be unrepresentative of efficient costs. In the case where X is significantly shorter than the historical sample period for the econometric models, and these models have a constant term and no time trend or time dummies, this adjustment acts to make some allowance for ongoing productivity and input price changes in the period covered by Y, relative to those which applied on average over the full historical sample period. Aside from decisions in these areas, adjustment is a mechanistic calculation.
Step 1C. Ongoing productivity (frontier shift) adjustment, applied from a year that appropriately reflects both model specification and choice of X in step 1B above.	○	○	●	○	Regulatory judgement informed by a range of evidence, such as estimates of the rates of productivity growth experienced in other UK sectors (e.g. figures derived from EU KLEMS dataset) and/or estimates of productivity growth experienced in the water industry. As indicated under projection method 2 in section 3 of this report, evidence concerning the cost trends of a notional efficient company

Projection method and key steps/adjustments involved	Factors captured by step				
	Cost driver variables	Notional efficient company	Ongoing productivity	Input price changes	Comments
					would be particularly relevant (rather than an industry-average company) – to help avoid double counting with step 1B. There is an argument that some estimates of the rate of sector-level productivity improvements from sources such as EU KLEMS already incorporate an element of catch-up efficiency improvements from less efficient companies within the sector, so it is possible that the rate of ongoing productivity improvement from a notional efficient company (if defined as something like an upper quartile performer) might be somewhat less than the sector-wide rate of productivity growth.
Step 1D. Conventional adjustment for input prices (e.g. RPEs), applied from a year that appropriately reflects both model specification and choice of X in step 1B above.	○	○	○	●	Regulatory judgement informed by a range of evidence, such as historical trends in - and/or external forecasts of - different categories of inputs used by water companies for retail activities. Ofwat's PR19 projection method for wholesale adopted a starting position that input prices grow in line with CPIH and only allowed adjustments for input price changes relative to this where it considered there was sufficient evidence of differences versus CPIH.
2. Application of assumed unit cost trend					
Step 2A. Modelled cost over forthcoming price control period estimated using econometric benchmarking models.	●	○	○	○	Modelled cost for forthcoming price control period draws on forecasts for explanatory variables and unit cost denominator over that period. Otherwise a mechanistic calculation based on estimation results from the suite of econometric models. If models have time trends variables these should not be applied beyond the end of the historical data period.
Step 2B. Conventional UQ-style adjustment based on comparisons of actual costs over last X years against modelled cost for that historical period.	○	●	○	○	Regulatory judgement on choice of notional efficient company (e.g. upper quartile or more/less demanding) and on historical time period over which to compare actual and modelled costs. Otherwise a mechanistic calculation. See further comments under step 1B above.
Step 2C. Adjustment for changes over time in unit costs for notional efficient company intended to reflect the net effects of productivity improvements and input price changes. Annual adjustment factor applied	○	○	●	●	Regulatory judgement, which could be informed by a range of evidence, including for example: (i) analysis of historical rates of changes in unit costs in water industry residential and non-residential retail activities; (ii) analysis based on results from econometric models of water industry residential retail costs over

Projection method and key steps/adjustments involved	Factors captured by step				Comments
	Cost driver variables	Notional efficient company	Ongoing productivity	Input price changes	
from a year that appropriately reflects both model specification and choice of X in step 2B above.					time; (iii) evidence of historical rates of change in unit cost metrics in other sectors of the UK economy; and (iv) evidence on possible adjustments for future period being different to that observed historically. As indicated under projection method 2 in section 3 of this report, evidence concerning the cost trends of a notional efficient company would be particularly relevant (rather than an industry-average company) – to help avoid double counting with step 2B.
3. Extrapolation from econometric models that involve a time trend					
Step 3A. Modelled cost over forthcoming price control period estimated using econometric benchmarking models with a time trend (time trend extrapolated in calculating modelled costs for that period)	●	○	●	●	Modelled cost for forthcoming price control period draws on forecasts for explanatory variables and unit cost denominator over that period. Otherwise a mechanistic calculation based on estimation results from the suite of econometric models. The time trend from the econometric model is intended to take account of the net impact of ongoing productivity and input price changes.
Step 3B. Notional efficient company adjustment based on comparisons of outturn costs over last X years of historical data versus modelled costs for the same period	○	●	○	○	Regulatory judgement on choice of notional efficient company (e.g. upper quartile or more/less demanding) and on historical time period over which to compare actual and modelled costs. Otherwise a mechanistic calculation. See further comments under step 1B above.
4. Forward-looking adjustment based on business plan comparisons					
Step 4A. Modelled cost over forthcoming price control period estimated using econometric benchmarking models with a constant term or time dummy variables (not a time trend)	●	○	○	○	Modelled cost for forthcoming price control period draws on forecasts for explanatory variables and unit cost denominator over that period. Otherwise a mechanistic calculation based on estimation results from the suite of econometric models.
Step 4B. Single forward-looking notional efficient company adjustment based on comparisons of business plan forecast costs over next price control period versus modelled costs over that period	○	●	●	●	Regulatory judgement on choice of notional efficient company (e.g. upper quartile or more/less demanding). Otherwise a mechanistic calculation. No separate adjustments for input prices and productivity applied in this case as these ought to be reflected in companies' forecasts

Projection method and key steps/adjustments involved	Factors captured by step				
	Cost driver variables	Notional efficient company	Ongoing productivity	Input price changes	Comments
					(NB this specific method does not apply if companies have been told to ignore effects of productivity and input prices when making the relevant forecasts). The calculation of the notional efficient company adjustment may be affected by differences in assumptions between Ofwat and water companies on forecast explanatory variables which feed into differences versus forecast and modelled costs.
5. Business plan cost forecasts included in the input data for the benchmarking models					
Step 5A. Using historical and forecast cost data in econometric models that have a time trend or time dummies and with modelled cost over forthcoming price control period calculated by applying the trend or dummies to that period.	●	○	●	●	Modelled cost for forthcoming price control period draws on forecasts for explanatory variables and unit cost denominator over that period. Otherwise a mechanistic calculation based on estimation results from the suite of econometric models. Essential that econometric model includes a time trend and/or time dummies rather than simply a constant term since a constant term would impose an unhelpful assumption that – other than for changes to cost drivers – retail unit costs are the same (e.g. in CPIH-adjusted term) over the historical data period and over the forthcoming price control period.
Step 5B. Notional efficient company adjustment based on comparison of business plan forecast costs over forthcoming price control period versus modelled costs for that period	○	●	○	○	Regulatory judgement on choice of notional efficient company (e.g. upper quartile or more/less demanding). Otherwise a mechanistic calculation. The calculation of the notional efficient company adjustment may be affected by differences in assumptions between Ofwat and water companies on forecast explanatory variables which feed into differences versus forecast and modelled costs.

Appendix 3: Supporting information for application of projection method 1

This appendix provides more detailed information on aspects of our approach to the application of projection method 1, which we summarised in section 3. The appendix takes the following topics in turn:

- Evidence on input price adjustments from PR19.
- Calculation of nominal input price adjustments for the period to 2030.
- Assumptions on ongoing productivity improvement the period to 2030.

Projection methods 2 and 3 are more straightforward to apply and we do not provide further information on these in this appendix.

Evidence on input price adjustments from PR19

In recent regulatory precedent, the analysis used for making adjustments for input price effects typically involves breaking down companies' costs into a series of separate categories and considering analysis and evidence to inform on estimated input price trends in each category.

Ofwat's approach to wholesale controls at PR19, which was followed by the CMA, adopted an approach involving various tests before making input price adjustments. This approach the effect of assuming input prices in a specific category grow in line with CPIH unless there is strong evidence to the contrary.

We collected some background information on the forecasts and assumptions on input prices from PR19. The table below shows some relevant figures from the PR19 process from Ofwat and the CMA, in respect of wholesale activities, and from Economic Insight in respect of residential retail activities in a report for Wessex Water. We distinguish between estimates on a CPIH-real basis and those in nominal terms.

Table 11 Examples of evidence and forecasts for input price changes from PR19

Source	Nature of estimate	Estimate (% change per year)	
		CPIH-real basis	Nominal basis
Ofwat PR19 FD	Allowed real price effect (RPE) adjustment for wholesale costs over period 2019/20 to 2024/25, based on labour costs but not for other cost types. Estimate based on wedge (1.2%) between the OBR forecasts of average earning growth and CPIH (except for last year, based on OBR forecast of labour productivity growth). Assumed labour cost share of 38.6% of total wholesale costs	Average over the period: 0.45%	

Source	Nature of estimate	Estimate (% change per year)	
		CPIH-real basis	Nominal basis
CMA determinations in the PR19 appeals (2021)	Allowed real price effect (RPE) adjustment for wholesale costs over period 2020/21 to 2024/25, based on labour costs but not for other cost types. Labour figure based on OBR forecasts for labour: 1.26% labour cost increase relative to CPIH applied to assumed share of labour costs within wholesale costs	0.48% on average for overall RPE adjustment	
Economic Insight report for Wessex Water (September 2018)	Estimate of gross input price effect for Wessex Water's retail costs over the period 2020/21 to 2024/25		Central case: 1.92%

The figure reported above from Economic Insight's work for Wessex Water in 2018 was built up from separate strands of analysis for different components of retail costs, with weights for each component based on the estimated contribution of these to Wessex Water's overall residential retail costs. In the table below, we summarise the implied average annual rates of change for each component (the Economic Insight report provides different figures for each year; we have condensed these to an annual average for simplicity).

Table 12 Calculation of annual average input price trends from Economic Insight (2018)

Retail cost component	Share of costs (at time of PR19 analysis)	Average annual rate of change estimated for 2020-25 (nominal)
Staff	38.07%	+2.2%
Doubtful debts	48.95%	+1.5%
IT	3.99%	+0.7%
Postage	2.87%	+6.7%
Other	6.12%	+2.0%

Source: Reckon analysis of Economic Insight (2018) *PR19 Retail Household IPP Analysis and Evidence A report for Wessex Water*, page 44.

In relation to the tables above, we highlight the following points:

- For wholesale activities, both Ofwat and the CMA focused their adjustments for input price effects on adjustments for labour costs. These used, in particular, OBR forecasts of wage increases, combined with an ex post adjustment mechanism based on ONS data from the annual survey of hours and earnings (ASHE). The allowances for labour costs were around 0.4% to 0.5% per year above CPIH inflation. Allowing for CPIH inflation, this is a little higher than the nominal labour cost inflation that Economic Insight estimated for retail activities in its work for Wessex Water at PR19.

- It might be thought that labour costs are more important to residential retail activities than to wholesale activities because, besides labour, the inputs needed for wholesale activities will include sizeable elements of construction and engineering materials as well as chemicals and energy. However, because bad debt costs make up such a large proportion of residential retail costs, the share of labour costs across wholesale and retail may not be that dissimilar. Indeed, in its analysis for Wessex Water at PR19, Economic Insight assumed a very similar share of labour within retail costs to that assumed for wholesale by Ofwat (38.1% versus 38.6%).

- In its work for Wessex Water at PR19, Economic Insight looked into the issue of postage costs, where there was evidence and analysis of these growing at a significantly faster rate than CPIH inflation (e.g. nominal increases of around 6.7% per year). However, on the figures reported by Economic Insight, postage was less than 3% of residential retail costs which means that the impact on input price inflation at the level of total retail costs would be quite limited.

- Economic Insight identified a cost category for IT costs but did not seem to find good data on input price trends for that category. The assumption used in this area looks less defensible than simply assuming that IT costs grow in line with CPIH.

Calculation of nominal input price adjustments for the period to 2030

Drawing in part on the methods and assumptions we saw from PR19, we then turned to consider how to make adjustments for nominal input prices for the 2025-30 period, in a way that was proportionate in the context of the wider project.

We calculated an index for nominal input prices for an efficient water retailer running from 2021/22 to 2029/30 using the following approach.

We split the costs of a notional water retailer into three main categories: bad debt costs; labour costs; residual costs. We treated these three categories as follows:

- We assumed for the purposes of this adjustment that nominal input prices affecting the bad debt component of residential retail costs was zero. This was to avoid double counting with the separate adjustment we made for differences between forecasts of the impacts on costs from difference between nominal and CPIH-real average bill changes over time.

- For the period 2021/22 to 2027/28, for which OBR forecasts were available, we assumed that labour costs would grow in line with OBR forecasts for nominal rate of changes in wages and salaries (this is an economy-wide forecast). For the period from 2027/28 to 2029/30, in the absence of OBR forecasts, we assumed that wages and salaries would grow at a nominal rate calculated as our forecast change in CPIH plus an uplift representing the average rate by which ONS outturn data on wages and salaries had exceeded CPIH growth in the period of our historical dataset (2013/14 to 2021/22).

- We assumed that residual retail costs would grow in line with CPIH forecasts (as explained earlier in this section, these were based on OBR forecasts in the period to 2027/28). These represented a small proportion of overall costs and it did not seem proportionate to look to further decompose these or identify alternatives to CPI/CPIH as the assumed input price trend.

We calculated the index by taking a weighted average of the assumed nominal input price growth rate for each of these three cost categories. Our approach to the weights was as follows:

- We assumed that in each year of the 2025-30 period, the ratio of (i) bad debt costs to (ii) the sum of labour costs and residual costs was equal to the ratio of bad debt costs to total retail costs in our calculation of modelled costs in that year. For this we calculated the ratio separately in each year for Bristol Water and Wessex Water, and for each of our model suites, and then took the average value.

- For the proportion of total costs which was not bad debt costs, we assumed a ratio between labour costs and residual costs in in each year of the 2025-30 period based on an Economic Insight report for Wessex Water at PR19 on the corresponding relative share of labour to residual costs.

Assumptions on ongoing productivity improvement the period to 2030

This subsection concerns the ongoing productivity within the context of projection method 1. The table below presents some estimates or assumptions for ongoing productivity at PR19, using the same sources of evidence from PR19 as for input price inflation above. A positive figure represents a productivity improvement and, in turn, a reduction in unit costs (all else equal).

Table 13 Examples of evidence and forecasts on ongoing productivity from PR19

Source	Nature of estimate	Value (% change per year)
Ofwat PR19 final determinations	Regulatory assumption on annual rate of ongoing productivity improvement (frontier shift) for wholesale water and wastewater activities	1.1%
CMA determinations in the PR19 appeals (2021)	Regulatory assumption on annual rate of ongoing productivity improvement (frontier shift) for wholesale water and wastewater activities	1.0%
Economic Insight report for Wessex Water (September 2018)	Estimated range for retail productivity savings in relation to opex (which is treated as most relevant to retail)	-0.42% to 1.1%

We highlight the following in respect of the figures above:

- For wholesale activities, the CMA used a 1.0% per year assumption for ongoing productivity (down slightly from Ofwat's figure of 1.1% from PR19 final determinations). A figure around 1% seems quite firmly entrenched in regulatory precedent. This figure was set in light of evidence on productivity in other UK sectors and, on face value, seems no less applicable to retail activities as to wholesale activities.

- A figure around 1% could be questioned in either direction, and the evidence base for this figure is hardly compelling. But we suspect that Ofwat and the CMA would need very strong evidence to set price controls on the assumption that ongoing productivity by a notional efficient company would be much less than 1% per year.

- We reviewed the section (section 4) on frontier shift productivity growth from the Economic Insight report for Wessex Water from September 2018. The central forecast from this report for the appropriate frontier shift productivity growth estimate for the 2020-25 price control period was 0.42% per year. We did not think that the report provided a strong case for this figure, which is much lower than the 1% assumption. For instance, the lower figure seems to reflect a significant reduction in the pace of UK productivity growth in more recent years, based on productivity figures for the UK as a whole or for large sectors of it. But it is not clear that the factors which may have contributed to this reduction (e.g. consequences of the global financial crisis) will apply to the same degree to the activities of monopoly water companies. A similar view was taken by the CMA in the context of the PR19 appeals.[22]
- There may well be an evidence base for a lower figure than 1% ongoing productivity growth, but we have not sought to explore this within the limited scope of this project (we agreed with the client companies to give greater focus for other areas of approach and assessment).

Aside from these points, we identified that a potentially important omission from the evidence base for ongoing productivity is evidence that is derived from data on the historical performance of water companies in relation tom their retail activities.

For the purposes of our cost projections, we decided to use two different assumptions on the productivity growth of a notional efficient company in the period from 2021/22 to 2029/30.

- A figure of 1% per year which fits with UK regulatory precedent above and some of the productivity evidence for the UK as a whole, and sectors within it, from EU KLEMS.
- An approximate estimate of the average annualised *productivity growth* of a notional efficient water retailer over the period from 2013/14 to 2021/22, drawing on (i) our estimates of the average annual change in unit costs between 2013/2014 and 2021/22 for a notional upper quartile retailer; and (ii) a high-level approximation for the average annual change in input prices relative to CPIH faced by such as retailer over this period. Our ongoing productivity assumption from this approach is a productivity improvement of 2.9% per year.

We described in section 3, in the context of projection method 2, our approach to estimating the change in unit costs between for a notional upper quartile retailer.

For the historical input price index used as part of the second estimate above, we adopted the following approach:

- We split the costs of a notional water retailer into three main categories: bad debt costs; labour costs; residual costs.
- We assumed that wages and salaries grew (relative to CPIH) in line with ONS data on annual changes in wages and salaries (this is an economy-wide figure).
- We assumed that residual retail costs grew in line with CPIH (i.e. a real price effect of zero).

---

[22] CMA (2021) *Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report*, paragraph 4.616

- We assumed that real input prices affecting the bad debt component of residential retail costs were zero (to avoid double counting with effects picked up via the average bill cost driver in the econometric models).
- We calculated the index by taking a weighted average of the assumed nominal input price growth rate for each of these three cost categories.

Our approach to the weights was as follows:

- We assumed that in each year of the historical data period, the ratio of (i) bad debt costs to (ii) the sum of labour costs and residual costs was equal to the average ratio of bad debt costs to total retail costs across the industry in that year.
- For the proportion of total costs which was not bad debt costs, we assumed a ratio between labour costs and residual costs in in each year of the historical period based on the Economic Insight report for Wessex Water at PR19 on the corresponding relative share of labour to residual costs.