IAP Response

Ref B2.8.WSH.CE.A1

# Cost Assessment Methodology

1 April 2019

## Contents

## 1. Overview

Our September business plan included costs of £3.5 billion for AMP7 to deliver our ambitious plan. Ofwat's initial assessment of plans provides their view of the efficient level of costs, which is some £640 million lower than our view. Our costs were robustly estimated and we have reconsidered the level of expenditure and we have largely kept the level of expenditure largely unchanged. The initial assessment of plans outlines Ofwat's approach to determining companies' cost allowance. We welcome the work that has been undertaken by Ofwat's cost assessment working group to work towards building a robust set of models for cost assessment. We have undertaken an initial review of Ofwat's approach to determining the 'efficient' level of expenditure and we have concerns in several elements of their approach. These areas are mostly around the use of Ofwat's models to determine the cost allowance. This chapter outlines some of our main areas of concern, as we currently see them in Ofwat's approach to modelling for the following areas:

- Ofwat's Approach to Econometric Modelling, including retail and wastewater growth;

- Ofwat's cost assessment framework, including the retails upper quartile adjustment and the interaction between cost and performance; and

- Ofwat's approach to frontier shift

## 2. Econometric Modelling

Ofwat's assessment of the efficient level of costs is made of several components of expenditure; base expenditure (botex), unmodelled based costs and enhancement expenditure. Ofwat has used different methodologies for determining the efficient level of costs for each of these cost elements. Base costs have been modelled using econometric models whereas Ofwat have used a combination of models and 'deep dives' for enhancement expenditure.

We welcome the work that has been undertaken by Ofwat's cost assessment working group to work towards building a robust set of models for botex cost assessment. We recognise the difficulties in modelling the base level of expenditure for seventeen very heterogeneous companies. The quality of the modelling depends on the quality and consistency of the data, both asset and cost data. The use of disaggregated modelling across the value chain requires consistent approaches on cost allocations between companies. Given that no model will perfectly account for all of these regional differences the results of the modelling need to be treated with caution. We have a number of detailed observations on both the base and enhancement models, however we highlight two main areas of concern; the approach to retail modelling and wastewater growth modelling

**Retail model Selection**

We welcome Ofwat's movement from an average cost to serve approach at PR14 to an econometric approach to determine cost allowances. The movement to modelling allows for 'environmental' factors that are beyond management control, including the level of deprivation, to be included within the assessment of costs. Whilst we welcome Ofwat's approach to econometric modelling we note that the models have a large range of residuals. In particular the bad debt models have a larger range of efficiency scores than the other models. The quality of the models depends on the amount of variation in costs that can be accounted by the explanatory variables but it is also dependent on the quality of the data. We note that the cost data used within the bad debt models is accounting data (rather than cost data) and there is a degree of subjectivity and potential inconsistencies between companies. For this reason, we think that horizontal audit work needs to be undertaken by Ofwat to investigate the consistency of the accounting cost data used in the retail cost models, in advance of the robustness of those models being assessed for their use at PR19.

Ofwat's approach to retail cost assessment triangulates nine econometric models. Ofwat's modelling technique accounts for several key cost drivers; a scale variable, deprivation, the average household bill, metering and transience. The measurement of deprivation is a key cost driver within the models, as approximately 45% of retail costs are bad debt and debt management. Ofwat considers three variables to account for deprivation; the percentage of households with default (Equifax), the council tax collection rate and percentage of households with default. Within Ofwat's modelling approach for retail we have several concerns:

- There is an inconsistency in the council data used in Ofwat's cost assessment between the data used in the econometrics (Data in RR1) and the forecast data (Data in RR3). Ofwat have acknowledged this difference in their query response and have noted that this will be updated for the draft determinations.

- Ofwat uses the percentage of households with default from Equifax as a proxy for deprivation. The data has been provided by United Utilities and procured from Equifax. Whilst data is available at a company level, we note that there is a significant lack of

transparency and the data cannot be replicated. The standard approach used by regulators, including Ofwat is to use assured company data or published government data where available.

- Ofwat uses the percentage of households with default variable (sourced from Equifax) in one of their bad debt models and two of the four total cost models (RTC1 and RTC2). This variable is insignificant in one of their total cost models (RTC1). As RTC1 has the same modelling specification as RTC2, except that RTC2 also accounts for economies of scale, it is unclear why Ofwat decided to include RTC1 in the final suite of models, given its poorer fit. We recommend that this model is removed from the suite of models.

**Wastewater growth modelling**

We have carefully reviewed Ofwat's approach to modelling Wastewater Growth expenditure. Ofwat have combined the costs for new developments and growth, growth at sewage treatment works and reducing risks of sewer flooding into one wastewater growth assessment, stating that these areas are interlinked with each other and are driven by population increase and demand growth. From our initial review we have identified two areas of concern:

- The inclusion of Hafren Dyfrdwy in the forecast modelling

- Ofwat's wastewater growth modelling does not adequately take account of the drivers of sewer flooding.

**Inclusion of Hafren Dyfrdwy**

Ofwat outlines several robustness tests that should be applied to their modelling for the PR19 econometric modelling consultation[1]. One of the tests is to consider the sensitivity of the results for the inclusion and exclusion of individual companies in the sample. Ofwat have excluded Thames and Southern from the metering enhancement model as these companies are outliers. Hafren Dyfrdwy and South Staffs and Cambridge have been excluded from the supply-demand balance enhancement model due to their outlying unit costs.

Wastewater growth expenditure is assessed using two regressions, firstly historical data and secondly using forecast data. The graph below plots the log of costs by the log of connections for the forecast data. There are several outliers in the bottom left quadrant relating to Hafren Dyfrdwy. The inclusion of Hafren Dyfrdwy significantly skews the regression, this is illustrated in figure 1 by the two different lines of best fit when including and excluding Hafren Dyfrdwy .

The impact of including Hafren Dyfrdwy is also illustrated by the significantly different assessment of allowed expenditure in the forecast growth model compared to the historical model. This anomalous result has arisen due to the significantly smaller size and particular operating characteristics of the substantially rural area served by Hafren Dyfrdwy compared to the other ten WaSCs. Therefore we believe the model should be reassessed excluding Hafren Dyfrdwy.

---

[1] CEPA (2018)- "PR19 Econometric Benchmarking Models"

**Dŵr Cymru Welsh Water**

Regression line including HD
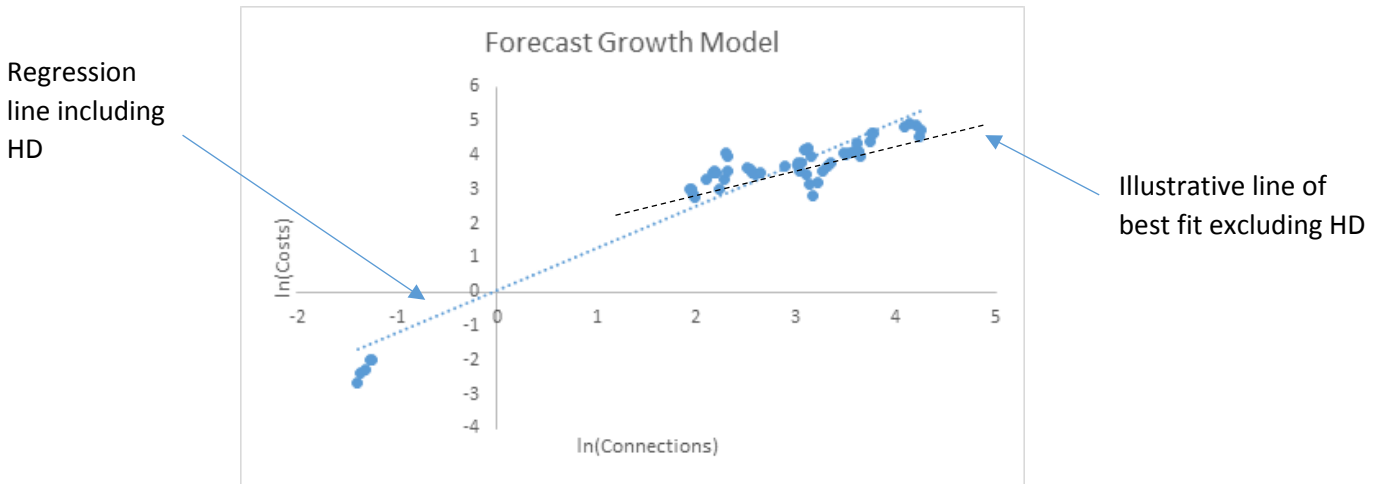
Illustrative line of best fit excluding HD

Figure 1: Forecast growth model results of DCWW assessment

**Sewer flooding modelling**

Ofwat has combined the costs for new developments and growth, growth at sewage treatment works and reducing risks of sewer flooding into their wastewater growth assessment. The efficient level of expenditure is assessed relative to the number of new connections. Whilst the number of new connections is an appropriate driver for new development, the model does not adequately take into account of the drivers of sewer flooding. There are three causes of sewer flooding from hydraulic overload; population growth, increased drainage area (urban creep) and climate change. Ofwat's modelling does not take account of urban creep nor climate change, which account for 98% of flooding from the hydraulic overload of our sewers. We believe this expenditure should be modelled separately or considered through a deep dive. Further evidence on the causes of sewer flooding are outlined in B2.2.22 PR19 Wastewater growth IAP response.

### 3. Cost Assessment Framework

One of the key elements to determine the efficient level of cost is the cost assessment framework. The cost assessment framework includes assumptions about the appropriate benchmark, frontier shift and how cost assessment fits in with the wider PR19 framework. We have concerns about several elements of the framework, especially the following two areas:

1. The use of a forward looking upper quartile and model selection in retail;

2. The separate assessment of upper quartile costs and service; and

**Forecast Upper Quartile Adjustment- Retail**

Ofwat's approach to residential retail modelling uses Corrected Ordinary Least Squares (COLS) to determining the 'efficient' level of expenditure. COLS is undertaken in two steps. Firstly, the cost function is estimated using an OLS regression. The estimated cost function is at the 'industry average' efficiency. The second step shifts the estimated cost function to the efficient frontier and may involve regulatory judgement.

A typical regulatory judgment involves an upper quartile adjustment based on the modelled data (i.e. first quartile of the distribution of efficiency scores over the modelled period). Ofwat's wholesale modelling and modelling at PR14 apply an upper quartile benchmark based on historical data. For retail expenditure, Ofwat deviates from the standard approach and determines the upper quartile adjustment by comparing the predicted expenditure from the model to companies' business plans forecasts over 2020/21 to 2024/25 (i.e. based on the 'out of sample' predictions and forecasts). In effect, Ofwat are shifting the OLS line estimated using data outside the estimated period. Ofwat note that they have used forecast business plan data instead of the historical data, as used in wholesale, as a result of the large reductions in the planned level of expenditure. However as the approach deviates from scientific and regulatory practices, we consider that it requires proper validation. For example, it is not clear whether Ofwat have tested if there is a structural break within the data to ensure that the historical data is consistent with the forward looking data. Oxera[2] have tested Ofwat's regression on the forecast data and note that their test results suggest that the forecast data is not supported by Ofwat's costs models. In effect, Ofwat's allowed costs are not supported by their own models. See Oxera's review of Ofwat's approach to residential retail cost assessment Appendix 1.

Ofgem considered using forward-looking benchmarks in their cost projections in RIIO-GD1 and RIIO-ED1. In their approach they tested whether the historical cost models supported the use of forecast data and where necessary made amendments to the models. At RIIO-GD1, Ofgem's cost predictions were based on both historical and forward looking upper quartile. However in their fast track assessment in RIIO-ED1, after model testing, Ofgem chose to consider only a historical benchmark.

Ofwat's approach to a forward looking upper quartile places significant weight on companies' business plan data. When setting the upper quartile target Ofwat should also take into account the impact of significant cost reductions on the level of service. There is a question mark on whether the cost reductions projected by companies will be achieved in practice. A comparison between business plan and outturn total retail expenditure for the first three years of PR14 control period (2016–18) shows that, on average, outturn costs is 2–6% higher than costs

---

[2] Oxera (2019)- "A review of Ofwat's approach to assessing residential retail business plan costs at IAP"

projected by companies in their business plans. This raises questions about the cost reductions predicted over AMP7. Ofwat has noted some reservations about using business plan data in their forecasts. Ofwat have noted that:

> "outturn costs often differ from companies' forecasts, which lead to movements in efficiencies and ranks…business plan rankings are not always indicative of actual efficiencies delivered, this is because companies will be incentivised to outperform the business plan assumptions and companies actual performance against their business plan will vary from company to company.[3]"

To determine the appropriate benchmarking for retail, Ofwat should test whether their models are supported by the forward looking data. Secondly careful consideration should be taken on the weight placed on companies' business plans level of expenditure. The strength of the upper quartile should be reflective of the level of ambition in companies' plans. Where companies have proposed large cost reductions there is a larger risk that these reductions may not be achieved. In considering an appropriate benchmark Ofwat should consider the weight that is applied to forecast data.

**Upper quartile of costs and upper quartile of service- Wholesale**

The initial assessment of plans sets cost allowances at the upper quartile efficiency of the industry. In Ofwat's outcomes framework, Ofwat requires companies to set targets for supply interruptions, pollution incidents and internal sewer flooding at the upper quartile level of performance in 2024/25.  Companies are also targeting a reduction in the level of leakage by a minimum of 15% and there is an expectation that companies should be setting stretching performance commitments towards the upper quartile on other measures. Ofwat state that the upper quartile level of performance for the three UQ performance measures and the 15% reduction of leakage should be achieved within base expenditure.

We have considered Ofwat's approach to setting cost and service targets and we have broken down their approach into two components:

1. The separate identification of upper quartile service and costs; and

2. The use of a forecast upper quartile service target (i.e. 2024/25 service level).

This section outlines the weaknesses in Ofwat's approach to separately identifying upper quartile costs and service and not taking into account any interactions. The graph shows the historical botex efficiency from Ofwat's water botex modelling and the 2016-17 water supply interruptions[4]. Ofwat sets the target level at the performance of the 25th percentile company, which for water is the 5th company. The graph shows the following:

- Only one company, Portsmouth Water has achieved both upper quartile efficiency and performance.

- The graph highlights a series of companies that are upper quartile on costs but performance is worse than the upper quartile.

---

[3] Ofwat (2015), 'Ofwat comments on Pennon Plc's initial submission to the Competition and Markets Authority (CMA)', July, p. 11.
[4] 2016-17 Data has been used instead of 2017-18 data due to the impact of the 'Beast from the East' and 'Storm Emma'.

- There are also a series of companies that are upper quartile on performance but have higher costs.
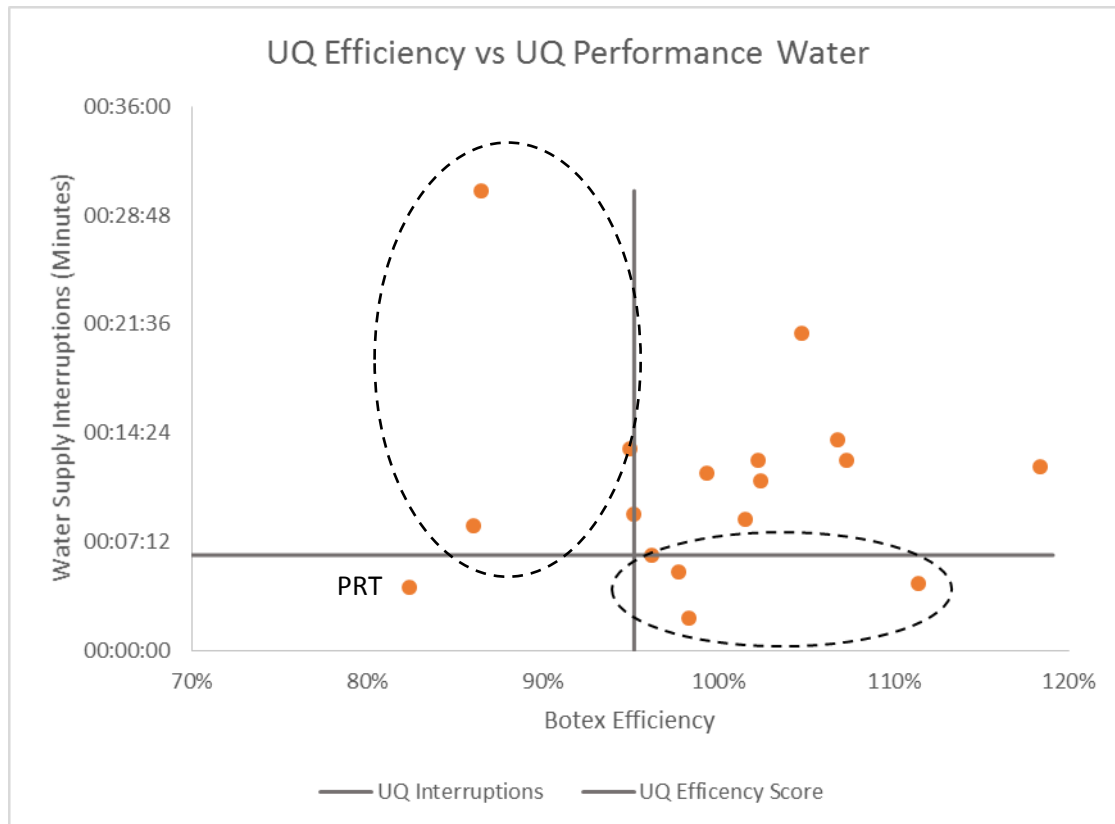


Ofwat's overall approach to UQ costs and service:

- Does not take into account the interactions between service and costs when setting upper quartile targets. There is a strong relationship between costs and service, these interactions should be accounted for as determining the upper quartiles separately can result in infeasible frontiers.

- Not only are the UQ cost companies not achieving current UQ performance, Ofwat are using a much more challenging forward looking upper quartile with no additional allowance for expenditure. The table shows the performance improvements from current performance to the 2024/25 upper quartile. There is a significant improvement in performance from the historical upper quartile to the forecast end of AMP target.

|  | 2017-18 Upper Quartile | 2024-25 Upper Quartile |
|---|---|---|
| Supply Interruptions | 6 minutes 18 Seconds[5] | 3 minutes |
| Pollution Incidents per 10,000km of Sewers | 25.6 | 15.1 |
| Internal Sewer flooding per 10,000km of Sewers | 1.7 | 1.28 |

- Alongside setting a stretching performance target, Ofwat has also applied a 1.5% per annum efficiency challenge on base costs. Overall Ofwat are expecting companies to achieve both the level of frontier shift and significant improvements in performance. The combined reduction in costs and improvement in service results in a double counting of productivity improvement by Ofwat (see following section).

- Ofwat's assessment also does not accounted for any historical expenditure for service improvements in the previous AMP as this would not be captured in botex. Any expenditure to deliver improvements in the level of service is classified as enhancement expenditure and therefore are not included within the base cost modelling.

- Companies have proposed target service levels in 2024/25, which Ofwat has used for its upper quartile target. Ofwat does not take into account any expenditure that has been included within companies' plans to achieve their target service level in 2024/25 and have also disallowed expenditure for service improvements in several areas.

---

[5] 2016-17 Data has been used for water supply interruptions to take out the effect of the 'Beast from the East' and 'Storm Emma'.

## 4. Frontier Shift

For wholesale, having arrived at figures for projected allowed costs in each year of AMP7, Ofwat then overlays a further 1.5% per annum real cost efficiency assumption for all companies to produce final figures for allowed botex. This comprises two elements: an assumption of the ongoing total factor productivity improvement that the sector could deliver in line with movements in the wider economy, and an additional element to represent the extra efficiency improvement that Ofwat thinks is achievable as a result of the introduction of the "totex and outcomes framework" at PR14. Ofwat implies that the first of the two accounts for about 1.0% of the overall frontier shift assumption, and the second accounts for 0.5%.

In addition, Ofwat considers whether there is a possibility that forecast real price effects relating to water company inputs, positive or negative, should be taken into account alongside the frontier shift projections. It concludes that no allowance for real price effects needs to be made.

Ofwat's proposals were informed by work that it commissioned from Europe Economics (the first element) and KPMG (both elements).

We do not think Ofwat's assumptions for either element have been properly derived or are realistic. We will make detailed representations in this area in the event that Ofwat has not modified them at the draft determination stage. These will include the following summary points.

In relation to the assumption of an economy-wide improvement in total factor productivity of 1% per annum:

- at a general level, the conclusion of Ofwat and its consultants that the economy at large and the water industry in particular can achieve 1% per annum total factor productivity gains is completely at odds with current evidence and the views of institutions such as the Bank of England and the Office for Budget Responsibility. The fact that productivity growth in the UK economy has been zero or negligible since the financial crisis a decade or so ago is the source of widespread puzzlement and concern. Ofwat offers no evidence as to why it should suddenly revert to pre-crisis growth rates at the start of AMP7;

- Ofwat's relatively bullish view on the prospects for productivity growth in the IAP for the purposes of projecting allowable costs stands in marked contrast to the assessment of wider macroeconomic trends that it presented in Appendix 12 to the final PR19 methodology paper published in December 2017. In chapter 5 – "Our Approach to the Cost of Equity" – the issue of stagnant productivity is cited as one of the principal justifications for Ofwat's choice of a much lower figure for the cost of equity than the estimate it used at PR14;

- in drawing on historical evidence and applying it to AMP7 projections, it appears that neither Ofwat nor its consultants has taken full account of the effect of replacing RPI as the basis for indexation in the sector with CPIH. Put simply, since there is a wedge of around 1% between RPI and CPIH, evidence that might support a 1% ongoing efficiency factor relative to RPI would not support a 1% ongoing efficiency factor relative to CPIH. Rather, the correct answer in this example would be zero. Or to put the point another way, Ofwat's proposal that companies "beat" CPIH by 1.5% is equivalent to projecting that companies can "beat"

RPI by 2.5%, which would be far more aggressive than any assumption made by a regulator in the UK in the past;

- there is an element of "double-counting" in the relationship between Ofwat's methodology on costs and its approach to service.  If, on average total factor productivity is improving by 1.5% per annum then, all else equal, a firm should be able to reduce its costs by 1.5% per annum holding service quality constant, or deliver better service whilst holding costs constant, or a blend of the two.  In taking the upper quartile of companies' *projected* service levels Ofwat has already effectively "used up" part of any future growth in total productivity growth, so it is not legitimate to also use all of it for cost improvements;

- overall productivity growth at the macroeconomic level is the average effect of firms that are innovating and forging ahead, firms that are catching up , firms that are standing still, and firms that are falling behind.  Since Ofwat's approach involves the application of significant "catch-up" *and* additional cost reductions in line with economy-wide total factor productivity improvements, it amounts to an expectation that the industry as a whole can improve productivity faster than the economy as a whole, but no justification for this is offered:  and this is before the industry-specific 0.5% improvement associated with the totex outcomes framework is added; and

- Ofwat makes no allowance for "real price effects" that would be expected to offset the effect of movements in total factor productivity.  We think this is a somewhat "extreme" position.  See John Earwaker's paper "A Review of Ofwat's PR19 Approach to Estimating Frontier Shift"[6]. We think Ofwat should revisit its approach to RPEs over the course of the summer, especially in light of a potential impact of Brexit.

In relation to the assumption of a further 0.5% per annum improvement associated with the introduction in 2015 of the "totex outcomes framework":

- the KPMG analysis on the basis of which Ofwat derives its assumption that a further 0.5% per annum can be achieved contains no evidence on efficiency trends in the sector since the new approach was introduced in 2015.  Rather, by its own admission, it offers a range of possibilities (0-1.2% per annum) based on an highly selective set of examples from other sectors, the relevance of which depends on critical and unproven assumptions, and which therefore needs to be interpreted with considerable caution; and

- in any event, since one of the main problems that the totex outcomes framework was intended to address was the alleged bias in spending decisions in favour of capital expenditure and to the detriment of operating expenditure solutions, the effect of the framework's introduction would be expected to involve more rather than less botex, all else equal.  It is counter-intuitive, therefore, to assume that its ongoing effect would be a reduction in botex in AMP 7, rather than an increase.

---

[6] Earwaker, J. (2019)- "A review of Ofwat's PR19 Approach to Estimating Frontier Shift" http://www.first-economics.com/PR19frontiershift.pdf

Appendix 1

A Review of Ofwat's approach to assessing residential retail business plan costs at IAP

oxera

# A review of Ofwat's approach to assessing residential retail business plan costs at IAP

Prepared for Dŵr Cymru cyf

March 2019

Final

## Executive summary

**Ofwat's use of a forward-looking benchmark in residential retail**

- Ofwat's approach of relying *solely* on a forward-looking upper quartile (UQ) benchmark to set cost allowances does not appear to be consistent with regulatory precedent and the validation process that is necessary to justify the use of forward-looking benchmarks.

- A key aspect of regulatory benchmarking exercises is deciding on the appropriate level of benchmark for the data that is modelled. For example, at RIIO-DG1 and RIIO-ED1,[1] Ofgem empirically tested whether the outturn-based cost models supported the forecast data, and made amendments in the model specification to enable this. Ofgem considered a combination of forward-looking and historical UQs to address a possible trade-off between the two—i.e. the achievability of a historical UQ and the ability of a forward-looking UQ to account for future step-changes in unit costs. A forward-looking UQ was not considered in isolation even where business plan data was modelled. It is not clear from the IAP publications whether Ofwat has considered a similar validation process. We note that, when we modelled forecast data separately or in combination with outturn data using Ofwat's outturn-based IAP models, the model outputs were economically/operationally counterintuitive and statistically weak.[2]

- Ofwat has not previously used a forward-looking UQ. In the PR14 wholesale and retail assessments and the PR19 wholesale assessment, it focused on historical data and a historical benchmark, as did the CMA in the Bristol Water price appeal inquiry.[3] Moreover, in recent merger inquiries, Ofwat has noted some reservations about using the efficiency ratios and ranking information based on business plan data.[4]

---

[1] Ofgem (2012), 'RIIO-GD1: Final Proposals - Supporting document - Cost efficiency', December; Ofgem (2014), 'RIIO-ED1: Final determinations for the slow-track electricity distribution companies', November.
[2] For example, we tested for a structural break in the relationship between cost and cost drivers over the forecast period; the test result suggests that the forecast data is not supported by Ofwat's cost models.
[3] Competition and Markets Authority (2015), 'Bristol Water plc, A reference under section 12(3)(a) of the Water Industry Act 1991, Report', October.
[4] Ofwat (2015), 'Ofwat comments on Pennon Plc's initial submission to the Competition and Markets Authority (CMA)', July, p. 11.

- Given significant cost reductions proposed by some companies over AMP7, relying purely on historical information could have drawbacks. At the same time, there is a question as to whether such cost reductions would actually be realised.

**Ofwat's IAP modelling approach in residential retail**

- Ofwat's retail models control for a limited number of cost drivers, as most of the variation in costs is expected to be driven by scale (costs expressed on a per-household basis). However, given the relatively limited set of drivers, their appropriate selection and ability to accommodate industry- and company-specific characteristics are key for prediction accuracy and reliability of results. Given this, we note that data issues with some of Ofwat's measures of deprivation, and its triangulation approach that gives equal weight to models despite differences in model quality, may lead to inappropriate allowances for Dŵr Cymru and possibly other companies.

- Additional work on the estimation approach and the accuracy of the retail econometric models may be necessary as part of further model development for the draft determinations, even if Ofwat were to use historical data and a historical benchmark. In the Bristol Water price appeal inquiry, the CMA felt that econometric models based on the high-level cost drivers that it considered and Ofwat's PR14 ones were likely to be susceptible to modelling limitations, and that a historical UQ benchmark might be overly demanding.[5]

# 1 Introduction

Dŵr Cymru has asked Oxera to review Ofwat's approach to assessing residential retail expenditure as set out in its initial assessment of business plans (IAP). We have been asked to assess the appropriateness of Ofwat's decision to rely *solely* on a forward-looking UQ benchmark in setting cost allowances. Our critique on this policy choice of Ofwat is presented in section 2 of the note. Our assignment also includes providing comments on other elements of Ofwat's modelling framework; in particular, the quality of Ofwat's IAP residential retail models. This is discussed in section 3 of the note.

# 2 Ofwat's use of a forward-looking benchmark

In contrast to its PR19 methodology in wholesale base expenditure, and its PR14 methodology on the wholesale and retail cost assessments, Ofwat estimates a forward-looking UQ efficiency challenge for residential retail expenditure. Ofwat has noted that this choice is driven by the decline in companies' projected costs over AMP7 relative to the current level of expenditure.[6]

Ofwat's decision to focus solely on the forward-looking UQ does not appear to be consistent with regulatory precedent and the validation process that is necessary to justify the use of forward-looking benchmarks. A key aspect of regulatory benchmarking practices is deciding on the appropriate level of benchmark for the data that is modelled. For example, when Ofgem considered the use of a forward-looking benchmark to account for companies' cost projections in RIIO-GD1 and RIIO-ED1, it tested whether the historical cost models supported the use of forecast data, making necessary amendments to the model specifications to enable this. In other words, cost models were also developed using forecast data (in isolation and/or in combination with historical data) to assess whether the efficiency ratios based on the business plan data (and thereby the forward-looking UQ) were appropriate for determining the allowances.

---

[5] The CMA therefore considered an average benchmark. Competition and Markets Authority (2015), 'Bristol Water plc, A reference under section 12(3)(a) of the Water Industry Act 1991, Report', October, paras 4.217–4.225.
[6] Ofwat (2019), 'Technical appendix 2: Securing cost efficiency', January, p. 21.

At RIIO-GD1, Ofgem rejected the models that used eight years of forecast data as they failed key diagnostics,[7] including a statistical test of whether the model coefficients (i.e. cost elasticities of the cost drivers) were stable over time.[8] Ofgem concluded that the immediate forecasts for the first two years were more robust, as the GDNs had made different assumptions in relation to some cost items.[9] Having undertaken necessary statistical validation of the forecast cost models, Ofgem set allowances using equally weighted historical and forward-looking UQ-corrected cost predictions.[10] At RIIO-ED1, Ofgem followed a similar model testing and validation procedure and chose to consider only a historical benchmark in the fast-track assessment, and a benchmark based on the entire historical and forecast period (13-year data) in the slow-track assessment.[11] As such, Ofgem's two-step approach to considering a forward-looking UQ consisted of (i) statistically evaluating the appropriateness of historical models on the forecast data; and (ii) using a combination of historical and forward-looking UQ based on the outcome of (i). It is also clear that, in RIIO-GD1 and RIIO-ED1, Ofgem did not focus solely on the forward-looking UQ, even where it was satisfied with the forecast data-based cost models.

Ofwat has not previously used the forward-looking UQ. In the PR14 wholesale and retail assessments and the PR19 wholesale assessment, it focused on historical data and used a historical benchmark, as did the CMA in the Bristol Water price inquiry.[12] Moreover, in recent merger inquiries, Ofwat has noted some reservations about using ranking information (and therein the efficiency ratios) based on business plan data. In particular, Ofwat has noted that:

> outturn costs often differ from companies' forecasts, which lead to movements in efficiencies and ranks…business plan rankings are not always indicative of actual efficiencies delivered, this is because companies will be incentivised to outperform the business plan assumptions and companies actual performance against their business plan will vary from company to company.[13]

From the technical annexes published as part of the IAP, it is not clear whether Ofwat has followed a validation process, as highlighted in the regulatory precedent, to ensure that the outturn-based models are robust when forecast data is included in them, or whether new models need to be developed to account for step-changes over the forecast period. For example, when we tested for structural breaks on Ofwat's IAP models and data, our results suggested that Ofwat's models might not be appropriate to pick up the relationship between costs and cost drivers over AMP7. Moreover, if the AMP7 data is modelled in tandem with historic data, most of the outturn-based models result in unintuitive and insignificant coefficients (see the regression results shown in the Appendix).

Clearly, there are advantages and disadvantages associated with relying on historical or forecast data. Historical data has the benefit of being anchored on actual outturn data. As such, the benchmark is based on actual observed

---

[7] Ofgem (2012), 'RIIO-GD1: Final Proposals - Supporting document - Cost efficiency', December, paras 4.2 and 4.9.

[8] Ofgem considered the F-test for parameter stability, which examines whether the regression coefficients are stable over time.

[9] Ofgem (2012), 'RIIO-GD1: Final Proposals - Supporting document - Cost efficiency', December, Appendix 4, paras 1.10 and 1.11.

[10] Ofgem (2012), 'RIIO-GD1: Final Proposals - Supporting document - Cost efficiency', December, Chapter 9.

[11] Ofgem (2014), 'RIIO-ED1: Final determinations for the slow-track electricity distribution companies Business plan expenditure assessment', November, paras 3.16 and 3.19.

[12] Competition and Markets Authority (2015), 'Bristol Water plc, A reference under section 12(3)(a) of the Water Industry Act 1991, Report', October.

[13] Ofwat (2015), 'Ofwat comments on Pennon Plc's initial submission to the Competition and Markets Authority (CMA)', July, p. 11.

performance and is thus achievable (to the extent that the benchmark is appropriate). However, historical data will exclude the potential for future innovations (or the impact of future increases in unit costs), so may be less (more) challenging unless a frontier shift assumption is overlain. In contrast, the provision and modelling of forecast data could reveal additional information that is unobtainable through historical data alone, such as the industry's views on how costs will change in the future.

In relation to business plan data, as per Ofwat's previous reservations about using efficiency estimates based on business plan data,[14] it is not clear if the cost reductions projected by companies will be achieved.[15,16]

We also note that, in light of the significant reductions in retail costs proposed by some companies over AMP7, relying purely on the historical information could have its drawbacks and will require careful overlay of frontier-shift and input price expectations. For example, if the historical UQ is used to adjust the forecast predictions (i.e. similar to the wholesale cost assessment framework), the underlying efficiency range will be wide, with the majority of the company projections coming out better than Ofwat's view. This suggests that an anticipated step-change in performance over AMP7 may need to be considered. However, for the reasons noted above, the forward-looking benchmark as currently applied is inconsistent with regulatory precedent on process, and additional evidential support is needed. To that end, Ofwat's current choice of basing its assessment solely on the business plan UQ—especially where, statistically, the historical models are not shown to be a good basis for predicting the future—will require a reassessment as part of its draft determinations.

## 3        Choice of propensity to default measure, model quality and estimation approach

In addition to relying on a forward-looking benchmark, Ofwat changed elements in its approach relative to the modelling consultation in March 2018.[17] These included the use of different cost drivers and a different estimation approach in the econometric assessment.[18]

As set out in the PR19 methodology document,[19] the use of econometric models in the IAP confirms Ofwat's move away from the average cost to serve (ACTS) approach used at PR14. This is an improvement in the assessment approach as it allows drivers that are not factored into the ACTS to be normalised for (e.g. deprivation, proportion of metered customers and economies of scale).

Ofwat's IAP models control for a limited number of cost drivers, as most of the variation in costs is expected to be driven by scale (costs are expressed on a

---

[14] Ofwat (2015), 'Ofwat comments on Pennon Plc's initial submission to the Competition and Markets Authority (CMA)', July, p. 11.

[15] The companies assessed at the forward-looking frontier have proposed to reduce unit retail total costs by 29% on average. Specifically, Southern reduced it by 50%, United Utilities by 27%, South Staffordshire Cambridge by 24%, Yorkshire Water by 15%, and South East Water by 11%.

[16] A comparison between business plan and outturn total retail expenditure for the first three years of the PR14 control period (2016–18) shows that, on average, outturn costs were 2–6% higher than costs projected by companies in their business plans. While cost performance over the last two years is needed in order to assess the divergence between business plan and outturn data over the last control period, it still raises questions; in particular, the impact on companies' revenues should the forward-looking benchmark deviate from the business plan.

[17] Ofwat (2018), 'Cost assessment for PR19: a consultation on econometric cost modelling', March.

[18] Random effects, RE, instead of OLS, ordinary least squares. See Ofwat (2019), 'Supplementary technical appendix: Econometric approach', January.

[19] Ofwat (2017), 'Delivering Water 2020: Our final methodology for the 2019 price review', December.

per-household basis). However, due to the relatively sparse model specification, the selection of appropriate cost drivers and their ability to accommodate industry- and company-specific characteristics are key for prediction accuracy and the reliability of results. In this respect, we note that Ofwat's measures of deprivation—a key driver of bad debt and total costs—may lead to inaccurate allowance estimates.

Ofwat considers three proxies for propensity to default: the **percentage of households with default**, the **council tax collection rate**, and the **percentage of household income deprivation**. The percentage of households with default variable (sourced from Equifax) is used in one of the bad debt models and two of the four total cost models (RTC1 and RTC2), in one of which (RTC1) it is not significant. As RTC1 has the same modelling specification as RTC2, except that RTC2 also accounts for economies of scale, it is unclear why Ofwat decided to include RTC1 in the final suite of models, given its poorer fit.

On the council tax collection rate variable, Ofwat's forecasts appear inconsistent with the historical data used for the modelling, which results in wide efficiency ranges. We understand that Ofwat has since acknowledged this error and noted that it would be addressed in the draft determination.[20]

More generally, Ofwat's disaggregated models appear to involve a higher degree of uncertainty than aggregate TOTEX models. In particular, bad debt models have larger efficiency ranges than the other suites. Other retail cost models have a low explanatory power, suggesting that the selected cost drivers do not explain much of the variation in the unit cost measure, as acknowledged by Ofwat in its technical annex.[21]

As part of further model development, it may therefore be necessary for Ofwat to undertake additional work on the robustness of the retail econometric models to determine the model specifications, triangulation approach and the appropriate benchmark. For example, in the Bristol Water price appeal inquiry, the CMA felt that econometric models based on high-level cost drivers and Ofwat's PR14 cost drivers are susceptible to modelling limitations and data errors, and that a UQ benchmark might be overly demanding.[22] The CMA also noted that, for it to properly apply a UQ (or any another benchmark besides the average), it would be necessary to make adjustments for company-specific factors to account for idiosyncrasies prior to calculating the efficiency scores.[23]

To assess the appropriateness of the benchmark, Ofwat could consider Stochastic Frontier Analysis (SFA).[24] As this estimation technique can separate noise from inefficiency at a company level, the efficiency scores estimated under SFA can be used to inform the choice of benchmark for a particular company.[25] This makes the choice of UQ (or another benchmark) less

---

[20] Ofwat (2019), Initial Assessment of Business Plans (IAP) queries, March, https://www.ofwat.gov.uk/wp-content/uploads/2019/02/Initial-Assessment-of-Business-Plans-March-2019.pdf.

[21] Ofwat (2019), 'Supplementary technical appendix: Econometric approach', January, pp. 34.

[22] Competition and Markets Authority (2015), 'Bristol Water plc, A reference under section 12(3)(a) of the Water Industry Act 1991, Report', October, paras 4.217–4.225.

[23] In the end, the CMA used an average benchmark based on outturn performance, and applied a cost trend reflecting productivity improvements and cost inflation. The CMA also observed that precedent from Ofgem and the Competition Commission shows that a less demanding benchmark than the upper quartile may be appropriate in cases where there is less confidence in the modelling results.

[24] Kumbhakar, S. and Knox Lovell, C.A. (2000), *Stochastic Frontier Analysis*, Cambridge University Press. Kumbhakar, S.C., Wang, H.-J. and Horncastle, A.P. (2015), *A Practitioner's Guide to Stochastic Frontier Analysis Using Stata*, Cambridge University Press.

[25] For example, if SFA predicts that a company will have a lower efficiency score than upper-quartile-corrected pooled OLS or RE, this may indicate that an upper-quartile benchmark is lenient and possibly

dependent on ad hoc adjustments and judgements. SFA could also be used to disentangle company-specific effects and inefficiency. The RE estimator, as currently used by Ofwat, assumes that the fixed, unobserved differences between companies (not accounted for by the cost drivers) are due to permanent differences in inefficiency, rather than uncontrollable differences in operating characteristics. As this assumption requires empirical validation, SFA can provide helpful guidance on what portion of such effects for individual companies can be attributed to inefficiency.

---

overcompensating for noise or advantageous company effects. The converse may hold if SFA predicts a higher efficiency score than pooled OLS or RE.

## A1    Regression results

The table below shows regression results when historic and business plan data are modelled together. The last row shows the p-value of the F-test used to test for structural breaks in the data. A p-value below 0.05 suggests that the null hypothesis of no structural break can be rejected.

|  | RDC1 | RDC2 | RDC3 | ROC1 | ROC2 | RTC1 | RTC2 | RTC3 | RTC4 |
|---|---|---|---|---|---|---|---|---|---|
| **Log (average bill size)** | 0.791*** | 1.031*** | 0.809*** |  |  | 0.502*** | 0.618*** | 0.600*** | 0.498*** |
| **Households with default (%)** | 0.097*** |  |  |  |  | -0.015 | -0.002 |  |  |
| **Council tax collection rate (%)** |  | 0.164*** |  |  |  |  |  |  | 0.079** |
| **Income deprivation (%)** |  |  | 0.049 |  |  |  |  | 0.002 |  |
| **Transience (%)** |  |  | -0.072*** |  |  |  |  |  |  |
| **Dual customers (%)** |  |  |  | 0.002 | 0.004*** |  |  |  |  |
| **Metered customers (%)** |  |  |  | -0.001 | -0.001 | -0.007** | -0.006** | -0.006** | -0.004* |
| **Log(connected households)** |  |  |  |  | -0.091** |  | -0.087** | -0.089** |  |
| **Constant** | -4.788*** | -19.634*** | -2.111*** | 2.833*** | 4.002*** | 1.258** | 1.437*** | 1.490*** | -7.043* |
| **Observations** | 200 | 200 | 200 | 200 | 200 | 200 | 200 | 200 | 200 |
| **Overall R-squared** | 0.659 | 0.579 | 0.650 | 0.0660 | 0.0925 | 0.604 | 0.640 | 0.635 | 0.643 |
| **F-Test (p-value)** | 1.93e-05 | 7.49e-09 | 0.00592 | 0.104 | 0.165 | 0.0290 | 0.0555 | 0.0158 | 0.0135 |

Source: Oxera.