WSX-C02 – Enhancement costs

Response to Ofwat's PR24 draft determination

Wessex Water
YTL GROUP

FOR YOU. FOR LIFE.

Representation reference: WSX-C02

Representation title: Enhancement costs

CONTENTS

1. Executive summary

Ofwat's overall approach to enhancement is comprised of three key approaches: (i) econometric benchmarking; (ii) deep dives; and (iii) shallow dives. We welcome Ofwat's objective to ensure all enhancement costs represent an efficient level, we would encourage a view that accounts for the considerations we set out below.

Overall approach

Between, and within, the three key approaches our enhancement cost areas are considered somewhat in isolation. Under a balanced approach, we would generally expect cross checks to ensure the median, or notionally efficient company is funded to meet all of its requirements. Without this the approach is somewhat limited. This is especially important in the context of a network industry where costs are complex and interrelated. We ask Ofwat consider this ahead of its final determination.

Approach to benchmarking

We support the use of cost benchmarking where it can be used to derive reliable estimates of efficient costs, and where the results are interpreted or cross-checked against all other relevant information. Given the complex and unique features of some investment proposals, we do not consider it is always an appropriate mechanism for setting cost allowances. Furthermore, we consider where benchmarking is used, the scope, specification, and results of models should be informed by - and cross checked against - engineering evidence.

Having reviewed Ofwat's cost models, there are a number of areas where there may be issues relating to the goodness of fit of models; and over-reliance on uncertain forecast data. Where the current approach is primarily relying on cost benchmarking for its enhancement cost assessment, we are concerned that Ofwat's use, and interpretation, of modelled outputs should recognise and account for the limitations of its models.

Our concerns on a model-by-model basis are set out in our individual enhancement cost representations.

Deep dive assessments

We welcome Ofwat's use of enhancement-specific engineering assessments. As discussed above, many enhancement schemes are unique by design, and this makes it very difficult to reliably compare or benchmark costs between companies. In these circumstances, it is important that one considers the evidence submitted as to why requested allowances are both necessary and efficient.

We submitted extensive evidence on our costs as part of our Business Plan. This included evidence on the need, scope, and efficient costs of these activities. In some areas, we do not consider that Ofwat has provided sufficient justification for the application and / or scale of efficiency challenges. Our response includes additional evidence which is aimed at addressing Ofwat's concerns before the final determination.

Shallow dives

To conduct its shallow dives, Ofwat has derived a company-specific challenge – capped at 20%. On this basis, the challenge for Wessex Water has been set at 20% (equivalent to £14 million). This represents a significant cut to our costs, which we do not believe appropriately reflects the efficiency of our costs. This cap has been set at twice the level it was at PR19, where it was 10%, with no explanation of the appropriateness of this change in practise. We suggest that final determinations should set shallow dive efficiency challenges at the minimum of: our updated efficiency score (which we expect to be considerably below 20%); or a cap of 10%.

There is also one enhancement area where we do not consider a shallow dive should be applied – **flow monitoring at STWs**. This investment area accounts for more than £10 million, which is above Ofwat's threshold for a shallow dive.

2. Approach to cost benchmarking

In principle, we support the use of cost benchmarking as a means of producing reliable estimates of efficient costs, and where the results are interpreted or cross-checked against all other relevant information. Given the complex and unique features of some investment proposals, we do not consider it is always an appropriate mechanism for setting cost allowances.

Having reviewed Ofwat's cost models, we are concerned that there are a number of areas where there may be issues relating to the goodness of fit of models; and over-reliance on uncertain forecast data. Where Ofwat is primarily relying on cost benchmarking for its enhancement cost assessment, we ask that Ofwat's use, and interpretation, of modelled outputs needs to recognise and account for the limitations of its models.

2.1. Model robustness

Econometric cost modelling is most appropriate as a tool for those enhancement programmes which are easily comparable across companies; involve a similar and repeatable activity; and have a clear cost driver.

Some enhancement programmes are highly company-specific. For instance, companies' WRC growth or P-removal programmes involve a number of discrete, bespoke schemes that all have their own individual characteristics depending on the location of each site, the nature of the upgrade (including any applicable permits), and other prevailing conditions such as the need for land purchase. As such, the costs of these individual schemes will have multiple cost drivers and their efficient unit cost (when normalised by a single metric relating to size) will often be very different. While some of these characteristics may 'average out' over a programme, this will not always be the case - for bespoke enhancement programmes such as WRC growth, IED upgrades or P-removal, the precise mix of schemes, and the extent to which they exhibit specific characteristics, will have a major impact on the efficient cost of that programme of works.

Ofwat has applied cost benchmarking to both these types of enhancement programme. But for the more complex programmes, for the reasons set out above, it is generally harder to fully explain efficient costs using simple benchmarking models. Multiple cost drivers can be added to capture a greater degree of the variation in cost, as Ofwat has done in some instances, but the usefulness of this approach relies on being to identify a sufficient number of explanatory variables that reflect the underlying drivers of cost. In these instances, we are concerned there is a risk that cost allowances derived from benchmarking models could overstate or understate true efficient cost allowances both at scheme level and at company level.

2.1.1. Goodness of fit

The adjusted R-squared value is a common measure of goodness of fit of econometric models. It captures the proportion of variation in the dependent variables (in this case, variation in costs) that is explained by the model. Ofwat has identified this as one of its key tests for assessing modelling robustness. It says that failure of this test

would raise serious concerns about using the model and 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".[1]

Table 1 summarises the adjusted R squared values from all the areas in which Ofwat has used cost benchmarking to set enhancement cost allowances. As shown, the models with the highest R squared (and are most successful at explaining variation in cost) are those enhancement activities which are more comparable and repeatable (e.g. smart metering, lead pipe replacement and supply interconnectors)[2]. In contrast, the goodness of fit of the models used to estimate costs for the most complex enhancement areas are low, with R squared figures well below 0.5 (highlighted in red in the table). Where an R-squared value is below 0.5, this means that less than half of the variation in costs from one scheme to the next can be adequately explained by the cost drivers in the model. This leaves a significant amount of unexplained or 'residual' variation, the scale of which is very unlikely to be explained by inefficiency (particularly when it is apparent at scheme level as well as company level).

Table 1 – Goodness of fit of Ofwat econometric models

Enhancement area	Range of adjusted R squared
Smart metering	0.952 to 0.961
Lead pipe replacement	0.912
Sanitary parameters	0.88
Supply interconnectors	0.7918 to 0.7921
P-removal (forecast data)	0.630 to 0.633
Storm overflows	0.596 to 0.666
STW growth	0.39 to 0.40
P-removal (historic data)	0.304 to 0.316
IED – secondary containment	0.201
IED – tank covering	0.078

The goodness of fit of Ofwat's chosen models can also be considered by looking at the range and dispersion of the implied efficiency scores calculated for each company. Looking at efficiency scores in this context provides another way to compare goodness of fit particularly in cases where models have different dependent variables, or where an R-squared value is not defined / applicable for the case (e.g. where Ofwat's cost benchmarking is a simple unit cost comparison). Efficiency score distribution is another of Ofwat's tests of model robustness that it uses to assess econometric models, noting that a large range of efficiency scores could indicate the presence of issues in the underlying model, such as the presence of omitted variables.[3]

Figure 1 below sets out this range for the same enhancement programmes as well as for Ofwat's other IED model. While there is a reasonably wide range for all areas, the range of scores is particularly high for Ofwat's IED models and its historic P-removal models, and (to a lesser extent) its company-level sanitary parameter model. This comparison provides strong evidence that there are particular issues with the modelling of these areas, and that there are likely to be important drivers of companies' efficient expenditure requirements that (i) differ between companies and (ii) are not captured by the drivers in Ofwat's models or unit cost comparisons.

[1] Ofwat (2024) *PR24 draft determinations: Expenditure allowances - Base cost modelling decision appendix*, page 73.
[2] For sanitary parameters, Ofwat used company-level rather than scheme-level benchmarking. We discuss the specific issues with this approach in our separate representation WSX-C09.
[3] Ofwat (2024) *PR24 draft determinations: Expenditure allowances - Base cost modelling decision appendix*, page 73.

Figure 1: Range of efficiency score for draft determination enhancement benchmarking

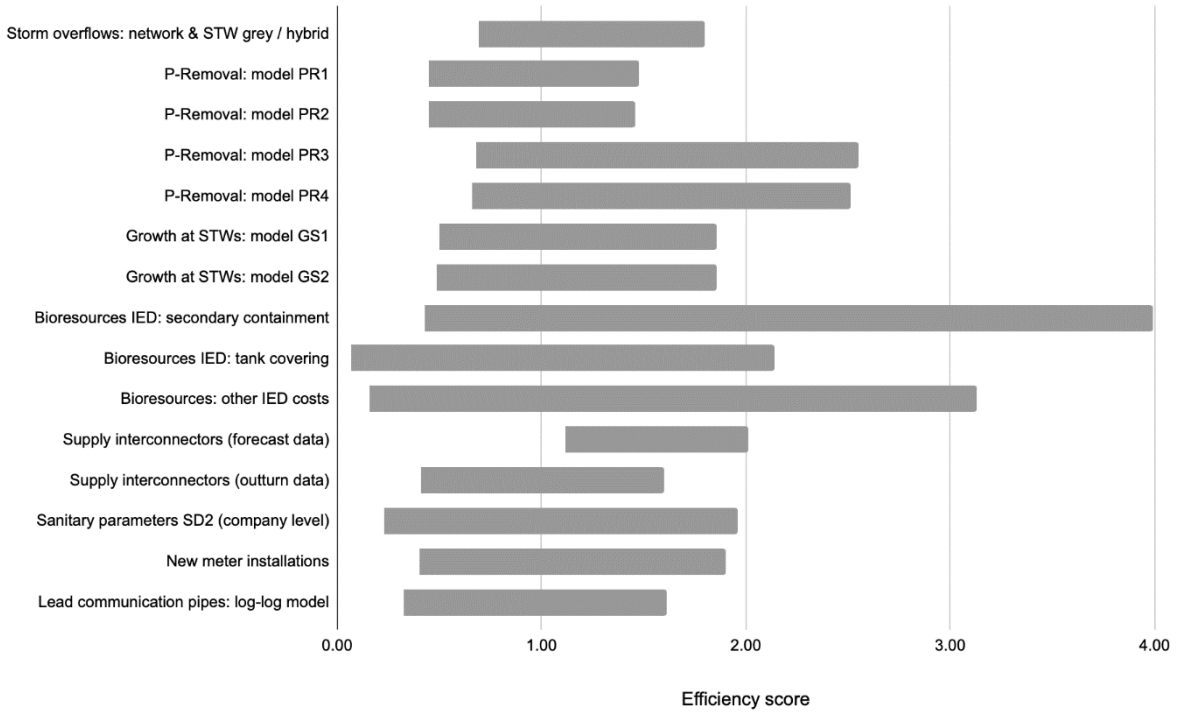
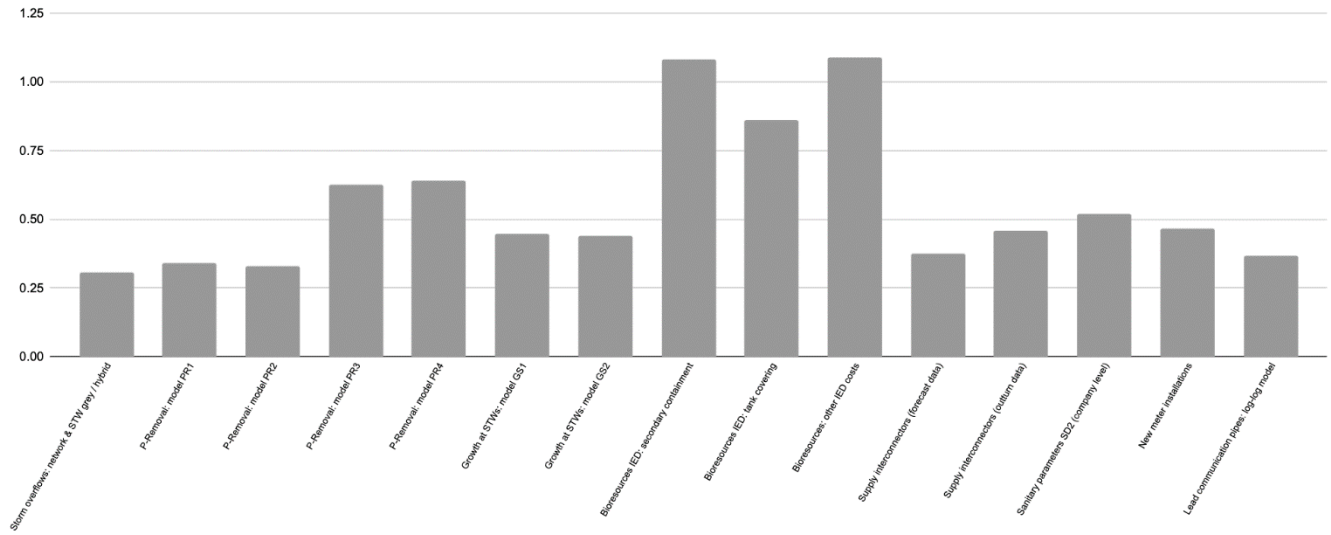


Figure 2 compares the standard deviation of efficiency scores across the same set of enhancement models as used for the chart of ranges above. This has the advantage of reflecting efficiency scores from the whole dataset rather than being driven of the maximum and minimum values which could be outliers compared to the rest of the data. Again, the standard deviation of efficiency scores for the IED and historic P-removal models programmes are considerably higher than that for each of the other enhancement models or categories from the chart.

Figure 2: Standard deviation of efficiency scores for draft determination enhancement benchmarking



Based on Ofwat's own tests, the results outlined above illustrate that there are issues with the predictive and forecasting power of its chosen models for some key enhancement areas. This can also be seen by considering the confidence intervals of some of the key parameters in Ofwat's modelling. For instance, for Ofwat's P-removal model, based on the reported statistics (see Table 2 below), the difference between Ofwat's modelled allowance and the upper bound of its 95% confidence interval for the constant is around £0.9 million (an average for PR1 and PR2 models), and around £0.5 million on average for all parameters. Given that this is a *scheme-level* estimate, this creates a material uncertainty range for the predicted costs at programme level, for a phosphorous removal programme of our size.[4]

Furthermore, scheme level estimates are directly used to set Price Control Deliverable (PCD) payment rates. This creates a separate risk that companies could be forced to return more than the true efficient cost of delivering a specific scheme, if some outputs are not completed by the end of AMP8.

Table 2: Model parameters and confidence intervals for smart metering

Area	Variable	Coefficient	95% confidence Interval	
P-removal (PR1)	pe_served	0.207	0.195	0.220
	consent_assumed_historical	0.219	0.025	0.412
	enhanced_consent	-4.069	-5.212	-2.926
	sq_enhanced_consent	0.803	0.496	1.109
	_cons	4.881	4.077	5.685
P-removal (PR2)	pe_served	0.206	0.194	0.218
	consent_assumed_historical	0.203	0.013	0.393
	enhanced_consent	-0.632	-1.055	-0.208
	p_below_025mgl	2.089	1.367	2.811
	_cons	2.403	1.423	3.384

Overall, the observed R-squared, standard deviation and confidence interval results indicate there is significant uncertainty regarding the appropriate cost allowances for quite a few enhancement programme areas where modelling is primarily used. Given the importance of ensuring that companies are appropriately funded for the efficient cost of their activities, many of which have statutory drivers, this is a cause for concern. We ask Ofwat to consider whether setting cost allowances for some enhancement areas based on models with such low statistical fit ensures that companies will be efficiently funded for their programme of works.

[4] By way of comparison, this range (for an individual phosphorous removal scheme) is larger than the fixed cost range implied by Ofwat's models for each company's total smart metering programme.

2.1.2. Outlier approach

Ofwat recognises that there is a need for separate treatment of outliers in its data. To do this, Ofwat applies Cook's distance statistic, which measures the influence of each observation on model estimation results. Ofwat used a threshold of four divided by the number of observations of the relevant model (4 / N) and dropped all observations with a Cook's distance metric that is higher than this threshold from the econometric models.

This is somewhat of a standard approach and Ofwat states this this approach ensures that the scheme level enhancement model estimation results are not affected by the inclusion of a limited number of influential observations. While this approach does help to lower the sensitivity of Ofwat's models to specific results, it does not fully address concerns relating to explanatory power. This is because the application of Cook's distance cannot, by definition, address any intrinsic limitations of any model. For example, where econometric models are being used inappropriately (such as to estimate costs of unique sites for which econometric models are not appropriate), application of Cook's distance will not materially change the fit of a model. This is a widely accepted principle.[5] We also note the adjusted R-squared values set out above are all reported after the application of Cook's distance, i.e. even on a sample of data without outliers, there is still significant unexplained variation in cost. All the issues set out in the previous section therefore continue to apply.

Furthermore, even when a model explains the relationship in the data reasonably well, Cook's distance will not necessarily identify all observations (or schemes) which are outliers in the sense that their costs are driven by factors not captured in the model. This is because it measures the influence of each observation on model estimation results, which is different, and will depend on other factors such as the size of the observation compared to the rest of the dataset. Cook's distance does not therefore guarantee that a given observation which require closer scrutiny will be identified.

In light of these issues, we ask Ofwat to take a pragmatic approach to the appropriateness of econometric modelling by price control area; and to consider the question of outliers using more than Cook's distance (i.e. taking a holistic approach to consider whether a model's explanatory power is strong enough; and considering engineering evidence as to whether the features of sites or schemes can be meaningfully compared). We have identified in our specific enhancement cost representation relevant instances where we consider that the specific circumstances and available evidence support a revised identification of specific outliers.

2.1.3. Summary

Overall, while we recognise that cost benchmarking can be an insightful tool for understanding relative efficiency, we ask Ofwat only rely on this approach in areas where there is sufficient evidence that cost models will reliably and robustly estimate the efficient cost of a given company's programme of works. There should be a high threshold for this; given the significant engineering evidence that we (and other companies) have provided in our business plan to set out the rationale for our proposed cost allowances in each area. Furthermore, we consider Ofwat should have regard for engineering evidence in all its cost assessments.

In areas where we consider that the (i) the engineering rationale clearly supports the use of cost benchmarking and (ii) the available evidence suggests a degree of confidence in modelling results, we are not therefore requesting any changes to Ofwat's proposed enhancement allowances in these areas[6]. This applies to smart metering and lead

[5] For example see lecture-20.pdf (cmu.edu) which sets out the following in relation to Cook's Distance: "Outliers are points that break a pattern. This can be because the points are bad, or because we made a bad guess about the pattern. […]. Deleting them, in order to make a linear model work better, would have been short-sighted at best."

[6] Though for smart metering we disagree with Ofwat's assessment of the base cost allocation and cost adjustment claim for smart metering costs. We discuss this in more detail in WSX-C07 (Enhancement costs – leakage and smart metering).

pipe replacement. This means that our smart metering allowance will be lower than that which was proposed in our business plan.

However, for a number of areas we are concerned that based on Ofwat's own tests, its models have limitations such that they cannot be reliably used to determine the efficient costs. For example, P-removal; IED enhancement and WRC growth. Furthermore, we are concerned Ofwat's approach to outliers does not address the weaknesses of these models and risks setting cost allowances which do not reflect the efficient cost of delivery.

These risks carry significant implications for companies' ability to deliver its enhancement programmes in AMP8:

- The scale of the distribution in efficiency scores implies very significant reallocations of costs. For example, Ofwat's PR1 p-removal model would result in reallocating £735 million in costs between companies – before any efficiency challenge is applied – compared to proposed costs in companies' business plans. We question whether the scale of this could be explained by relative efficiency assumptions in plans and under / over-forecasting on the part of companies.

- As noted above, Ofwat's scheme-level benchmarking is also being used to set PCD payments, as well as cost allowances. While some residual cost variation may average out for companies at programme level (or even between programmes), the application of these results to PCD payments means that companies' delivery will be tracked against the scheme-specific estimates from these models, and could be compelled to return more than the actual cost of delivering a scheme. We consider this necessitates an even higher level of confidence in the accuracy of cost modelling results than in previous price reviews.

For these reasons, we ask Ofwat to consider its approach ahead of the final determination. We have set out in our individual enhancement cost representations how we propose that Ofwat amends its approach where relevant, to take account of these issues.

2.2. Use of forecast data

Ofwat's econometric models generally rely on the use of company forecast data. In principle, these should reflect companies' best *ex-ante* view as to a stretching yet deliverable settlement. This is particularly the case for enhancement investments, which are often entirely new (or at least very different) programmes of work. Forecast data underpins the actual enhancement programmes that companies are seeking to deliver in AMP8 for their cost allowances, and is generally more appropriate than 'backward-looking' models based on historical data that may not represent the new kind of work being undertaken.

However, we ask Ofwat to consider the risks associated with this, and ensure appropriate cross-checks are undertaken. This is because companies themselves are not subject to perfect information, and the outturn evidence suggests forecasts allowances at PR19 may have been too optimistic, as set out below.

To consider the use of forecast data, we have reviewed the degree to which companies' PR19 forecasts have been borne out in outturns[7]. This shows that:

- Across wholesale services, the forecasts of base expenditure for AMP7 which companies submitted in their business plans at PR19 are projected to be exceeded by **9%** on average in AMP7. Only three out of 16 water

[7] Annex 1 to this document sets out our supporting analysis in respect of retail cost allowances.

and wastewater companies are projected to spend below what they forecast, with the remaining companies all projected to spend more than forecast (some by over 20%)[8].

- For retail services, this picture is even more stark. There is an average overspend by companies of **15%** over the AMP7 period, compared to their PR19 forecasts.

This overspend has not translated into better performance on outcomes. As shown in our separate representation WSX-R01, companies have largely incurred net penalties across AMP7 and underperformed by -0.65% (relative to regulatory equity) across the four years of AMP7.

Furthermore, this overspend has been more pronounced among companies who were most ambitious in their PR19 plans. Ofwat's efficiency challenge for retail in AMP7 was heavily informed by the upper quartile (UQ) of companies' forward looking efficiency ratios, calculated as the ratio of companies' business plan forecasts to modelled costs. For these companies, residential retail costs are projected to exceed the PR19 forecasts by 26%. The projected overspend for other companies is 10.2%. If projected overspend over AMP7 for these companies was simply in line with average projected overspend for non-UQ companies (not even in line with actual outturns), this retail efficiency challenge would be 14.4% over the AMP instead of 20.6%. This represents a substantially lower forward-looking efficiency challenge than what Ofwat derived based on the less accurate, upper quartile forecasts.

A similar pattern can be seen in respect of wholesale operating costs – particularly for wholesale water, where UQ companies are projected to have an overspend on their operating expenditure of 27.8% compared to 22.3% for non-UQ companies. For wastewater, the projected overspend is 16.2% and 11.9% respectively.

This indicates not just that companies have understated their required efficient cost allowances, but that companies who have been most ambitious in forecasting efficient stretch have generally overspent the resulting cost allowance by more than others. This is somewhat unsurprising given Ofwat's approach provides companies with an incentive to demonstrate high *ex-ante* ambition when it comes to efficiency. We have previously raised concerns that these incentives, and Ofwat's broader approach to base cost assessment, may encourage companies to submit expenditure forecasts in their business plans that are lower than what they realistically expect to incur over a price control period. This concern is borne out by the comparisons set out above.

We also consider that the choice of benchmark for an efficiency adjustment is affected by the effectiveness of the underlying models. In the presence of omitted variables, an adjustment based on the upper quartile efficiency score can reflect an overestimation of the efficiency of the companies with an upper quartile efficiency score. This is because a company may have costs different to the median (i.e. at the UQ) because of model error (e.g. misspecification), rather than because of any inherent efficiency. Given the likelihood of omitted variables in several of Ofwat's enhancement benchmarking models (discussed above), this is a further reason why an UQ efficiency factor could be mis-calibrated.

For these reasons, we consider Ofwat should refrain from applying a more stringent upper-quartile efficiency challenge. In its Draft Determination, Ofwat has proposed doing this in two areas – for storm overflow STW solutions; and for IED secondary containment and other costs. However, Ofwat has said that where it has applied an average efficiency benchmark, it will review whether this provides a sufficiently stretching catch-up efficiency

[8] Moreover, the extent to which companies are projected to exceed their PR19 forecast is considerably greater for operating expenditure than it is for capital maintenance. This indicates that companies may have been controlling expenditure by postponing investment decisions and stretching asset lives beyond design lives, which is possible (at least in the short term) for capital maintenance and renewals but not for operating expenditure. To the extent this has been the case, it means the observed overspend in AMP7 base expenditure compared to forecasts is understated.

challenge at final determinations. We consider the uncertainty that is present in the PR24 models, as well as the available evidence from PR19 and AMP7, clearly demonstrates that a median benchmark is more appropriate.

Overall, we consider this indicates the need for caution in (i) interpreting forecast data in general; and (ii) setting industry-wide efficiency challenges based purely on one or two companies' forecasts. We ask that Ofwat considers this when setting its final determination, particularly in light of the greater level uncertainty at PR24 which warrants further scrutiny on the achievability of forecast costs.

Finally, we also consider that where there is considerable uncertainty regarding costs, scope or requirements that Ofwat consider the introduction of an uncertainty mechanism framework. As part of our response to Ofwat's Draft Determination, we have proposed two types of uncertainty mechanism. The inclusion of these mechanisms in the PR24 settlement would help to reduce the information asymmetry that exists not just between companies and the regulator but also between companies and their own supply chains, by allowing aspects of the PR24 determination to be set when there is much greater certainty over efficient cost allowances. In doing so, it would help to mitigate the risks and impacts set out above – though it is important to note this would not remove the need to fully scrutinise and challenge forecasts, particularly those which are most ambitious, for the parts of the AMP8 settlement that will continue be determined *ex-ante*. In this sense, it is a complementary solution which we believe will support the industry to deliver the AMP8 programme more efficiently for customers.

3. Deep dive assessments

Where Ofwat considers that an enhancement investment area does not lend itself to statistical modelling or unit cost benchmarking, but the requested investment is material, Ofwat has based its cost assessment on the written evidence provided in company business plans. It does this through a 'deep dive' process where business cases are assessed against three criteria:

- **Need for investment**: evidence that the proposed enhancement is required, including alignment with agreed strategic planning frameworks where relevant. Where a need is not fully evidenced, it applies a cost challenge of between 0% and 100% based on its assessment.
- **Best option for customers**: evidence that an appropriate number of options has been considered, benefits quantified, and cost benefit analysis and best value assessment undertaken, accounting for customer views where appropriate. Where Ofwat accepts the need for the investment, but companies have not provided sufficient and convincing evidence that the investment is the best option for customers, it applies an efficiency challenge ranging from 10% for minor concerns, 20% for some concerns and 30% for significant concerns.
- **Cost efficiency**: evidence that expenditure requests should be based on efficient costs with sufficient and convincing evidence to demonstrate efficiency. Where Ofwat considers this evidence to be lacking, it again applies an efficiency challenge ranging from 10% for minor concerns to 30% for significant concerns.

We welcome Ofwat's use of enhancement-specific assessments. As discussed above, many enhancement schemes are unique by design, and this makes it very difficult to reliably compare or benchmark costs between companies. In these circumstances, it is important that Ofwat considers the evidence that companies have submitted in their business plans as to why requested allowances are both necessary and efficient.

We submitted extensive evidence on our enhancement costs as part of our business plan. This included evidence on the need, scope and efficient costs of these activities. In reviewing this evidence, we do not consider that Ofwat has always provided sufficient justification for the application. Furthermore, we are concerned that the scale of cuts made under the framework risks underfunding (i.e. that the % reductions are high, and somewhat arbitrarily set). As set out in WSX-C04, Ofwat's approach to assessing what is funded in base is directly contradictory to its approach to assessing what is funded in base costs specifically. For example, in relation to enhancement Ofwat considers

companies are able to make the appropriate trade-offs to prioritise (e.g. in relation to our proposal to significantly increase CCTV monitoring); however, in relation to base costs Ofwat consider companies are funded to deliver a specific level of output (e.g. in relation to metering or mains replacement). It is our view that both cannot be simultaneously true.

Our specific concerns with Ofwat's approach are set out in our individual enhancement cost representations.

In order to work constructively with the regulator, our response includes additional evidence which is aimed at addressing Ofwat's concerns. This is provided in our individual representations on enhancement costs where relevant.

Taking account of this evidence, we therefore expect to see considerable movement in the final determination.

4. Shallow dive assessment

For less material enhancement investment lines (defined as expenditure which is less than 0.5% of the water or wastewater wholesale totex, and less than £10 million), Ofwat has carried out a shallow dive assessment. Ofwat has derived a company-specific challenge by examining the cost efficiency of companies in the enhancement benchmarking models used in our cost assessment – capped at 20%. Ofwat has set our shallow dive efficiency challenge at the capped value, i.e. 20%.

We are concerned that the level of the cap represents a stretching efficiency target for costs which have not been examined in detail. Furthermore, we note this is a significant increase from PR19 where the equivalent cap was set at 10%, which Ofwat said balanced the risk of excessive disallowance of costs[9]. A 20% cost challenge is highly significant and will prevent the delivery of the full programme of activities in the areas where a shallow dive approach has been used. For instance:

- A 20% reduction has been applied to a number of Catchment Partnership projects focussed on delivering habitat and water quality improvements under the wastewater price control. Examples of these include the Resilient Avon Project in the Hampshire Avon and the Chew Catchment Project. This reduction also impacts the level of match funding leverage from wider partners and their ability to deliver. Professor Penny Johnes, the University of Bristol academic lead for the Chew Valley Partnership Project (08MU100852), has highlighted: *"these [in-kind or financial] offers [from catchment partners] were proportionate to the Wessex Water input to each investigation, so there would not just be a 20% reduction in WW expenditure, but a matching reduction in contributions from the extensive stakeholder networks you have built and curated for each catchment."* In the Chew Valley Partnership Project, we had identified a need for £2 million in AMP8 funding for this WINEP deliverable with commitments of £6.6m from the wider partners, which will be proportionally reduced as a result of the 20% cost challenge.

- Our contribution to the water industry collaboration to deliver the Chemical Investigations Programme will also be reduced. A number of these WINEP actions relate to projects being delivered in collaboration with other companies where financial contributions are proportional to the population served. However, two WINEP actions have been included specific to Wessex Water which will further the sector's understanding on specific issues: 08WW100028a - 4f innovative pathway control investigation; and 08WW100053a - 4g Local investigation Ratfyn WRC. These investigations are complex, requiring bespoke academic research and cannot be managed by

[9] PR19-draft-determinations-Securing-cost-efficiency-technical-appendix.pdf (ofwat.gov.uk).

UKWIR. Reducing the funding allowance by 20% will have an impact on the supporting academic institutions reducing the level or emerging contaminant analyses and ecotoxity impact assessments, respectively.

These two examples relate to Partnership projects where there are efficiencies in enabling investigations and implementation delivery at catchment scale. These Partnership projects mean that greater outcomes can be delivered for the environment and customers than if Wessex Water were to attempt to deliver these alone and very much echoes the expectations outlined in the Government's Strategic Policy Statement to Ofwat:

"[We expect Ofwat to] encourage companies to operate in partnerships across catchments maximising co-funding and green finance opportunities, wherever appropriate, including through market mechanisms."

In other areas, a 20% efficiency cut will have similar implications for the deliverability of the overall programme.

Change requested

We ask that Ofwat sets it shall dive assessment efficiency challenge at the minimum of the following values:

- **Our updated efficiency score.** If Ofwat corrects its enhancement cost assessment to reflect the points raised in this representation and our other enhancement cost representations, our modelled efficiency score will be substantially less than 20% for water and wastewater.
- **An efficiency score capped at 10%** as per Ofwat's approach at PR19. While we disagree with Ofwat's overall approach to shallow dives (for the reasons set out above), we consider this is a constructive and pragmatic solution which works with the regulator to achieve a draft determination which is acceptable.

There is also one enhancement area where we do not consider a shallow dive should be applied – **flow monitoring at STWs**. This investment area accounts for more than £10 million i.e. above Ofwat's threshold for applying a shallow dive. Furthermore:

- Ofwat's own assessment indicates that our unit cost for flow monitors is below the industry median and that under a modelled cost approach we would be awarded *more* than our requested allowance. Ofwat's own assessment sets out that our flow monitoring programme is "overall efficient compared to unit cost benchmark assessment." In this context, Ofwat has not explained why it considers a 20% efficiency challenge is appropriate to make.
- Secondly, we have revised our flow monitoring costs based on more recent actual outturn data for flow monitoring installations. We consider this demonstrates clearly why our revised proposed cost allowance represents an efficient programme of installations.

Ofwat has said that it has applied discretion to carry out a deep dive assessment for some areas below (but close to) the shallow dive threshold, however it does not explain why this area has been treated as a shallow dive when it exceeds Ofwat's stated threshold.

For these reasons, we request that Ofwat undertakes a fuller assessment of our flow monitoring costs rather than applying a shallow dive challenge. We have provided further evidence in our separate representation WSX-C09 as to why our costs for this enhancement area are necessary and efficient.

5. Responses to Ofwat's Draft Determination consultation questions

For completeness, we set out in Table 3 below our responses to Ofwat's specific Draft Determination questions on enhancement cost allowances, or where they can be found.

Table 3 – Ofwat Draft Determination proforma – enhancement cost actions

Question	Response
Do you agree with our decision to use OLS to estimate our scheme level enhancement models?	We support the use of cost benchmarking where it can be used to derive reliable estimates of efficient costs, and where the results are interpreted or cross-checked against all other relevant information. Where models are used, simple OLS models have merit as they are transparent and can generally be easily interpreted. However, given the complex and unique features of some investment proposals, we do not consider this is always an appropriate mechanism for setting cost allowances. This is reflected, among other things, in the adjusted R squared and efficiency score distributions associated with Ofwat's OLS models for these areas – particularly p-removal and IED. Ofwat's tests for assessing model robustness rightly identify that a low R squared and a high efficiency score range gives causes for concern that an OLS model has low explanatory power and will not accurately predict efficient costs for a given scheme (and a company-level programme). Our full reasoning for this is set out in this representation and in other enhancement cost representations.
Do you agree with our decision to exclude outliers based on a Cook's distance threshold of 4 / N?	While we agree that is important to consider the presence and impact of outliers in any cost benchmarking exercise, we do not consider that applying Cook's distance statistic will address this issue. The application of Cook's distance cannot address any intrinsic limitations of any model. Furthermore, even when a model explains the relationship in the data reasonably well, Cook's distance will not necessarily identify all observations (or schemes) which are outliers in the sense that their costs are driven by factors not captured in the model. In light of these issues, we ask Ofwat to take a pragmatic approach to the appropriateness of econometric modelling by price control area; and to consider the question of outliers using more than Cook's distance (i.e. taking a holistic approach to consider whether a model's explanatory power is strong enough; and considering engineering evidence as to whether the features of sites or schemes can be meaningfully compared). We have identified in our specific enhancement cost representation relevant instances where we consider that the specific circumstances and available evidence support a revised identification of specific outliers.

Question	Response
Do you agree with our approach to setting an efficient enhancement expenditure allowance for outlier schemes?	We agree that a deep dive assessment should be used in these circumstances. As discussed above, many enhancement schemes are unique by design, and this makes it very difficult to reliably compare or benchmark costs between companies. In these circumstances, it is important that Ofwat considers the full evidence that companies have submitted in their business plans as to why requested allowances are both necessary and efficient, and to allow companies' full requests where the evidence demonstrates that requested costs are necessary and efficient. However, we do not consider that Ofwat has always provided sufficient justification for the application of its deep dive assessments. We have provided further evidence where relevant to ensure Ofwat's assessment. We also note that Ofwat's outlier approach is 'one-sided' in the sense that companies are awarded a portion of the difference between modelled and requested allowances when modelled allowances are lower, but not vice versa. Ofwat should consider this in the context of its overall benchmarking approach and striking the correct balance between the need to secure cost efficiency, while ensuring that companies are adequately funded for to deliver their IED enhancement programme.
Do you agree with our decision to apply the PR19 log-bias adjustment to address log-bias (where relevant)?	We agree that this adjustment is necessary. As Ofwat notes, log-bias is a downward bias in model predicted costs for the sector, which would lead to allowances being set too low. All other things equal, it is therefore appropriate to correct for this.
Do you agree with our decision to set the efficiency benchmark at the company level instead of scheme level?	We broadly agree that a company-level approach is preferable to a scheme-level approach. However, for the reasons set out in this representation, we disagree with the use of an upper-quartile efficiency challenge. We consider the uncertainty that is present in the PR24 models, as well as the available evidence from PR19 and AMP7, clearly demonstrates that a median benchmark is more appropriate. Our full views on this are set out in Section 2.2 of this representation.
Do you agree with our approach to assessing grey and grey-hybrid storage storm overflow enhancement costs?	We consider Ofwat's approach to setting allowances for grey and grey-hybrid solutions is reasonable, though we are requesting that Ofwat updates its modelling to reflect the changes to our storm overflow investment requirements for 2025 to 2030. As set our above, we also disagree with Ofwat's application of an upper-quartile efficiency challenge for network solutions.
Do you agree with our approach to assessing phosphorus removal enhancement costs?	We disagree with the proposed approach to setting phosphorous removal cost allowances. In summary, we consider that the complex and unique features of phosphorous removal programmes mean that a benchmarking model cannot be reliably used to predict efficient expenditure requirements in this area, particularly at scheme level. This is borne out by the results of the modelling, and is consistent with engineering rationale. We also disagree with the weight placed on models derived using historic costs which are unlikely to reflect the forward-looking efficient costs of companies' phosphorous programmes. Our full reasoning is set out in WSX-C09 (Enhancement costs – wastewater treatment).

Question	Response
Do you agree with our approach to addressing the implementation issues associated with modelling phosphorus removal enhancement costs?	We have no specific views on Ofwat's post-modelling adjustment, though we note that they do not address the more fundamental issues covered above.
Do you agree with our approach to identifying overlap with base costs so that customers do not pay for non-compliance with existing permits?	All our growth at STW schemes meet Ofwat's criteria for providing allowances for growth in AMP8. We consider this also applies to new schemes that we have included in our updated plan. This is discussed in more detail in WSX-C10 (Enhancement costs – Water Recycling Centre (WRC) growth).
Do you agree with the models we have selected to explain differences in efficient growth at STWs enhancement costs?	We agree that the choice of functional form and selected choice of cost drivers is reasonable. However, we consider that there is still significant variation in scheme-level costs that is not being explained by the modelling approach. This creates a major risk that Ofwat's cost modelling will not capture all the factors that determine scheme-level efficient costs and, depending on the mix of programme schemes, could lead to companies being underfunded for the efficient costs at programme level. We have set out how this can be addressed, particularly through the treatment of specific outliers. This is set out in Section 2 and Section 4.2 of WSX-C10 (Enhancement costs – Water Recycling Centre (WRC) growth).
Do you agree with our approach to adjusting modelled allowances to account for costs incurred outside of the 2025-30 period?	We broadly agree that this is a reasonable approach to take.
Do you agree with our approach to adjusting allowance to account for past under-delivery?	We disagree with Ofwat's approach. The reasons for this are set out in Section 4.4 of WSX-C10 (Enhancement costs – Water Recycling Centre (WRC) growth).
Do you agree with our approach to setting efficient IED secondary containment, tank covering and other IED cost allowances	We disagree with the proposed approach to setting IED enhancement cost allowances. In summary, we consider that the complex and unique features of IED requirements mean that a simple benchmarking model or a unit cost comparison cannot be reliably used to predict efficient expenditure requirements in this area, particularly at scheme level. This is borne out by the results of the modelling, and is consistent with engineering rationale. Our full reasoning is set out in WSX-C18 (Enhancement costs – bioresources and IED).
Do you agree with our approach to assessing sanitary parameters enhancement costs	We disagree with the proposed approach to assessing sanitary parameter enhancement cost allowances. In summary, we consider that Ofwat has not taken sufficient account of the scheme-level data. Our full reasoning is set out in WSX-C09 (Enhancement costs – wastewater treatment).
Do you agree with our approach to addressing the implementation issues associated with modelling sanitary parameters enhancement costs?	See our answer above.

Question	Response
Do you agree with our approach to assessing new meter installation and meter upgrade costs?	We respond to this question in WSX-C07 (Enhancement costs – leakage and smart metering).
Do you agree with our decision to assess smart infrastructure costs within the meter installation and meter upgrades models?	We respond to this question in WSX-C07 (Enhancement costs – leakage and smart metering).
Do you agree that the number of lead communication pipes replaced or relined is the key factor that explains differences in efficient costs?	We agree that number of lead communication pipes replaced or relined is the key cost driver for companies' lead pipe programmes.
Do you agree with our approach to triangulating between the median unit cost and an econometric model?	We consider that an econometric model (adjusted for log bias) is generally preferable to a simple median unit cost as it is able to better capture where there is a non-linear relationship between the costs and the size of the overall programme (i.e. a cost-volume elasticity less than one).
We welcome views on other model estimation methods we could consider using to estimate scheme level econometric models given that we have multiple schemes for each company but no time dimension.	We have no specific views on this question. However, for certain enhancement areas such as P-removal and IED which are highly company-specific, we consider it is unlikely that applying a different model estimation method or applying a different specification could fully explain all the material efficient drivers of costs, and are therefore unlikely to be able to reliably predict efficient expenditure requirement. In these and other areas, we believe that Ofwat's use, and interpretation, of modelled outputs should recognise and account for the limitations of its models. We have set out in our individual enhancement cost representations how we propose that Ofwat amends its approach to take account of this, reflecting the specific circumstances of each enhancement area.
We welcome views on the approach to setting the efficiency benchmark in our scheme level enhancement cost assessment.	See our response to the question above.

Annex 1 – Use of forecast data to set cost allowances

We have sought to understand the degree to which Ofwat should place reliance on companies' expenditure forecasts when it reviews and challenges plans and sets a forward-looking cost allowance.

We clearly will not know how companies' AMP8 expenditure forecasts (as set out in their PR24 business plans) will pan out until after the close of that price control period. We can, however, look back at AMP7 to understand how the expenditure forecasts which companies had set out for AMP7 in their PR19 business plans compare with the expenditure that companies have actually incurred in the years to-date, and with the more recent and updated forecasts for the expenditure in the two remaining years of the AMP, 2023-24 and 2024-25.

In this annex, we outline the approach we have taken to doing this for retail analysis, before presenting the results. Our approach to assessing wholesale costs followed the same approach.

A1-1. Approach

Our analysis rests on comparing the expenditure that companies had forecast for AMP7 as set out in their PR19 business plans, with the expenditure that companies have incurred over the three years to-date in that AMP – from 2020/21 to 2022/23 – and their recent forecasts for the remaining two years, 2023/24 and 2024/25.

Our analysis was based on the following main steps:

1. We compiled data on companies' forecast of their retail expenditure for AMP7 based on their PR19 business plans, specifically on the version of plans which Ofwat drew on for final determinations.
2. We compiled data on companies' outturn retail expenditure in AMP7 – covering 2020/21, 2021/22 and 2022/23 – and augmented it with companies' recent forecasts, as set out in their PR24 business plans, for the two remaining years of AMP7.
3. Drawing on the above, we calculated the difference between companies' AMP7 projected retail expenditure and the forecasts of such expenditure which they had put forward in their PR19 business plans.
4. We examined the variation across companies in the gap between PR19 forecasts and projected retail expenditure. In doing so, we distinguished between companies, in particular, in terms of:
    - Whether the forecast expenditure set out in their PR19 placed them within, or without, a forward-looking upper-quartile (UQ) efficiency ratio. We derived this UQ ratio by considering the ratio of each company's forecast expenditure for AMP7 from the PR19 business plan to its modelled expenditure for AMP7. We calculated the latter by running the set of cost assessment econometric models Ofwat used at PR19
    - The difference, in percentage terms, of the PR19 forecast retail expenditure for AMP7 compared to historical spend over the five-year period from 2015/16 to 2018/19.

A1-2. Findings

Figure 3 presents the industry-wide over/under-spend of PR19 residential retail forecasts in AMP7. Companies in orange are those that were identified as forward-looking UQ at PR19.

Figure 3 - Residential retail: AMP7 projected over/under-spend compared to PR19 forecast

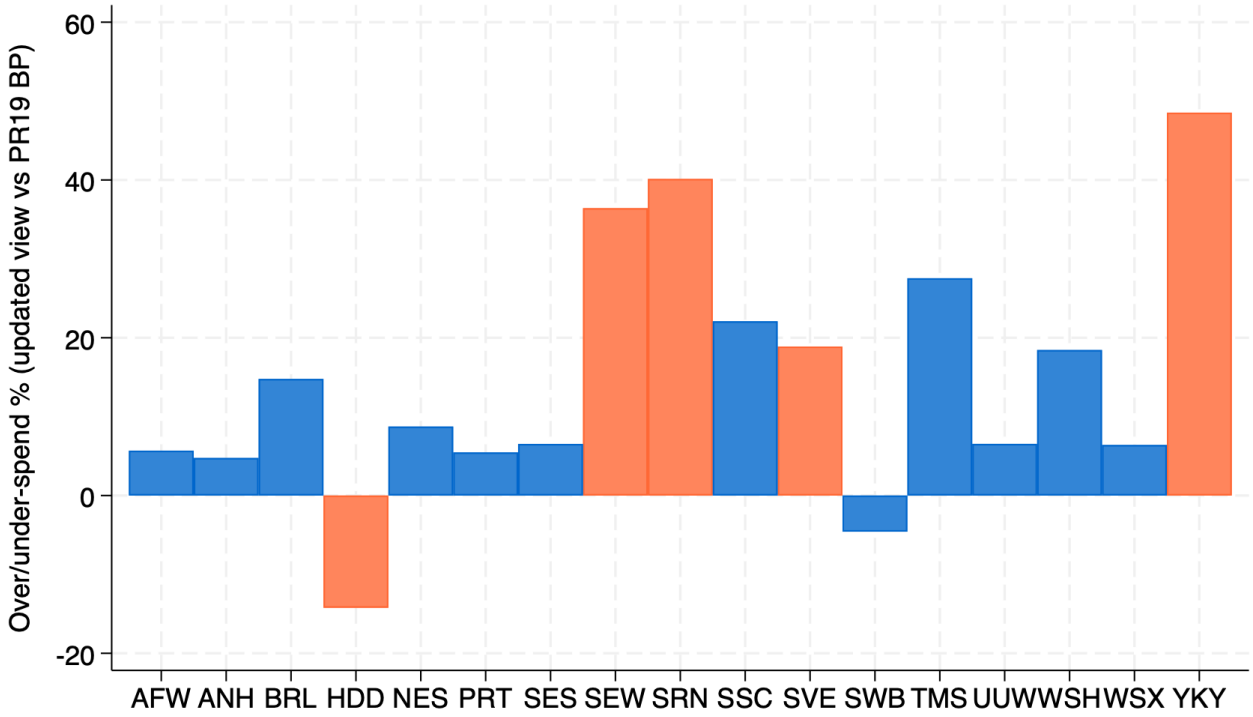


Figure 3 illustrates a number of points. Firstly, there is a clear overspend at industry-level; all but two companies overspent relative to their forecasts, and across all companies the average overspend was 14.8%. It is perhaps not surprising that companies have over-spent their PR19 retail forecasts as a whole, given the impact of Covid-19 at the start of the AMP; and the impact of unexpectedly high inflation on nominal costs, which are the figures used in this analysis, in line with Ofwat's approach at PR19.

Secondly, this overspend was much greater on average among UQ companies (26.0%) compared to non-UQ companies (10.2%). The margin of this additional overspend is substantial (over 15 percentage points). Moreover, of the upper-quartile companies, only Hafren Dyfrdwy under-spent its PR19 forecast. If we exclude HDD, the upper-quartile overspend figure would be 36.0%, 25.8 percentage points higher than the non-UQ company average.

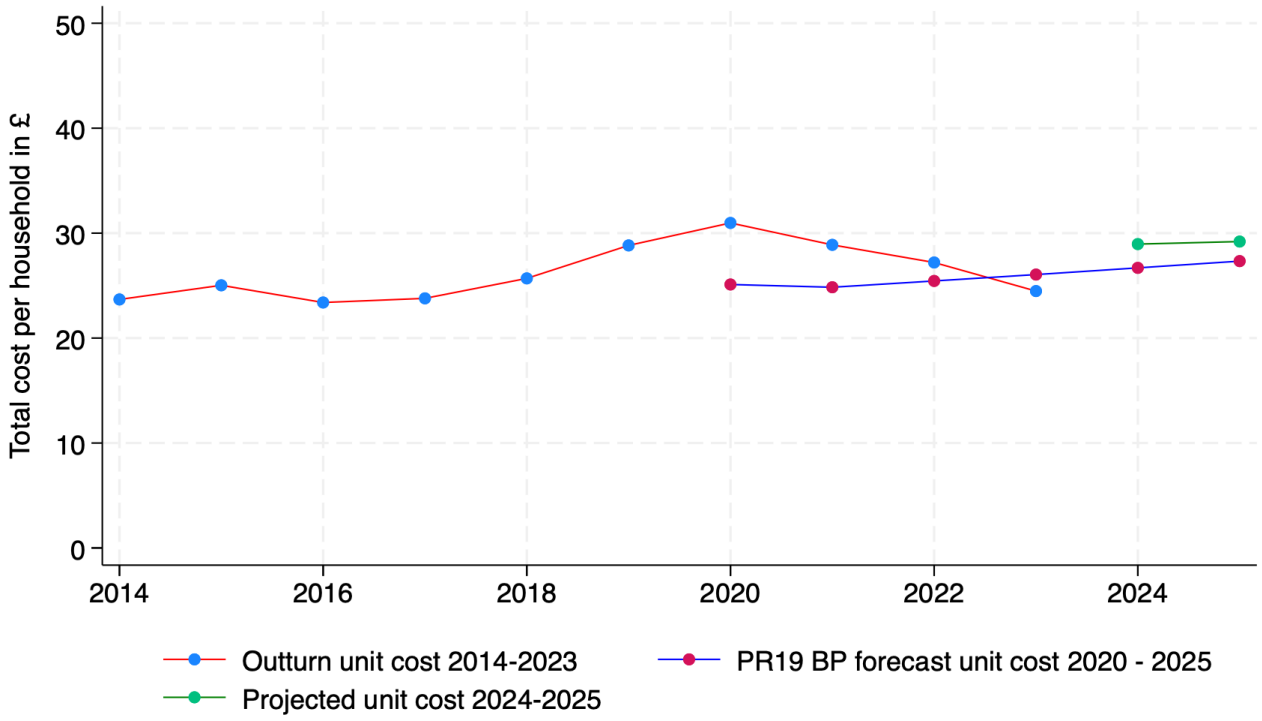
We have calculated what the forward looking upper-quartile efficiency ratio would have been, had the cost forecasts of those companies which were identified at PR19 as UQ been such that the average overspend of UQ companies been in line with the rest of the industry. We have done this by adjusting the PR19 forecasts of each of the upper-quartile companies whilst holding the outturn figures constant, so that their calculated overspend was in line with the non-upper quartile average overspend. For example, if a given UQ-company over-spent its forecast by 20% whilst the non-UQ average overspend was 12%, we calculated the forecast the UQ company would have had to submit at PR19 such that its overspend would be calculated as 12%. From this we then recalculated an adjusted view of upper-quartile, simulating what the efficient level of costs would have been if the UQ-companies had submitted forecasts with a similar level of inaccuracy to the rest of the industry. Our finding from that exercise is that the forward-looking upper-quartile efficiency challenge would have been 14.4% rather than the 21.6% which results from using the forecast figures that companies actually submitted in their plans.

A1-2.1. Company-level analysis

To further illustrate this point, we have directly compared our PR19 outturns and forecasts with two companies (Southern Water and Yorkshire Water) whose PR19 forecast costs placed it within the forward-looking UQ set of companies.

Figure 4 below presents outturn actual expenditure figures for Wessex Water from 2013/14 – 2022/23, and compares this with forecast expenditure figures provided at PR19 (the purple line) and updated forecasts for the last 2 years of AMP7 (the green line). This shows that while we overspent relative to our PR19 forecast for the first three years of AMP7, primarily due to the effects of Covid-19, our 2022/23 outturn cost figure is below the forecast made at PR19. PR24 forecasts for the remaining 2 years of AMP7 are in line with the initial forecast. This shows that while costs were higher than anticipated early in the AMP, updated costs to the end of AMP7 are largely in line with PR19 forecasts.

Figure 4 - Wessex Water AMP7 residential retail unit costs



Figures 5 and 6 present the same comparison for Southern Water and Yorkshire Water respectively. The lines in this case tell a different story. At PR19, Southern Water indicated an expectation of a sharp decrease in costs for the first 2 years of the AMP, followed by a gradual tapering to finish substantially lower by the end of the AMP. As with Wessex Water, costs in the first years of the AMP were higher than forecast, however unlike Wessex Water, Southern Water has been unable to bring costs down to the levels forecast at PR19. Furthermore, the updated projection for the final two years of AMP7 indicate a notable increase in costs relative to the forecast made at PR19.

Figure 5 Southern Water AMP7 residential retail unit costs

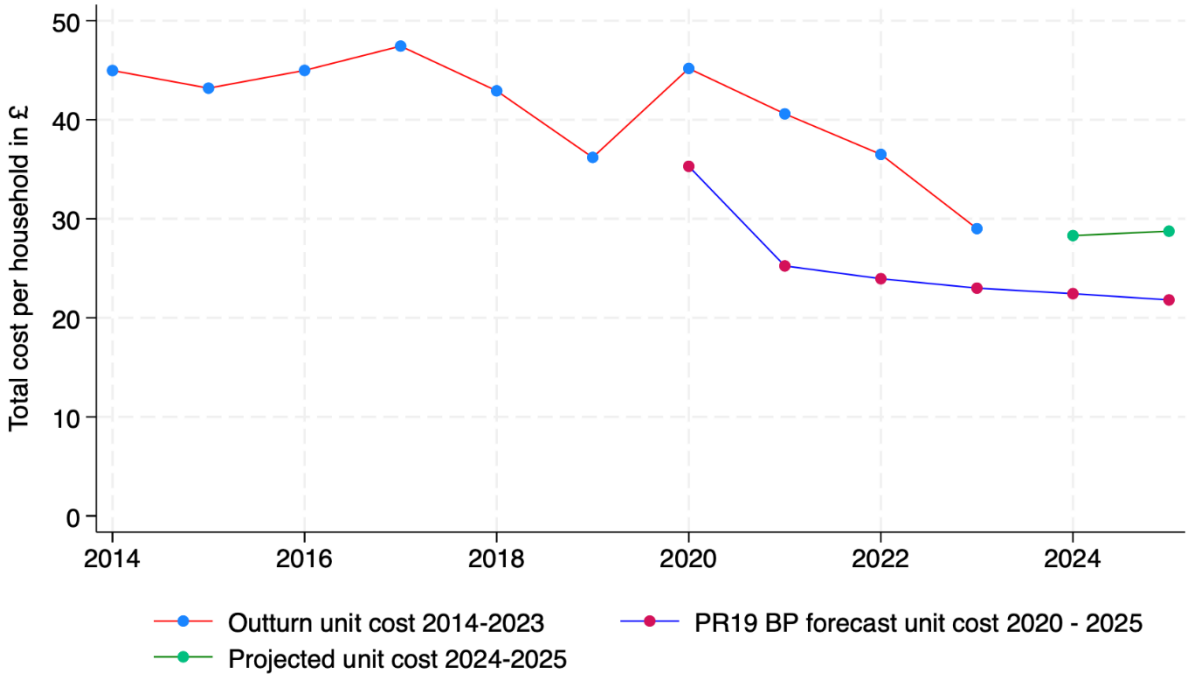
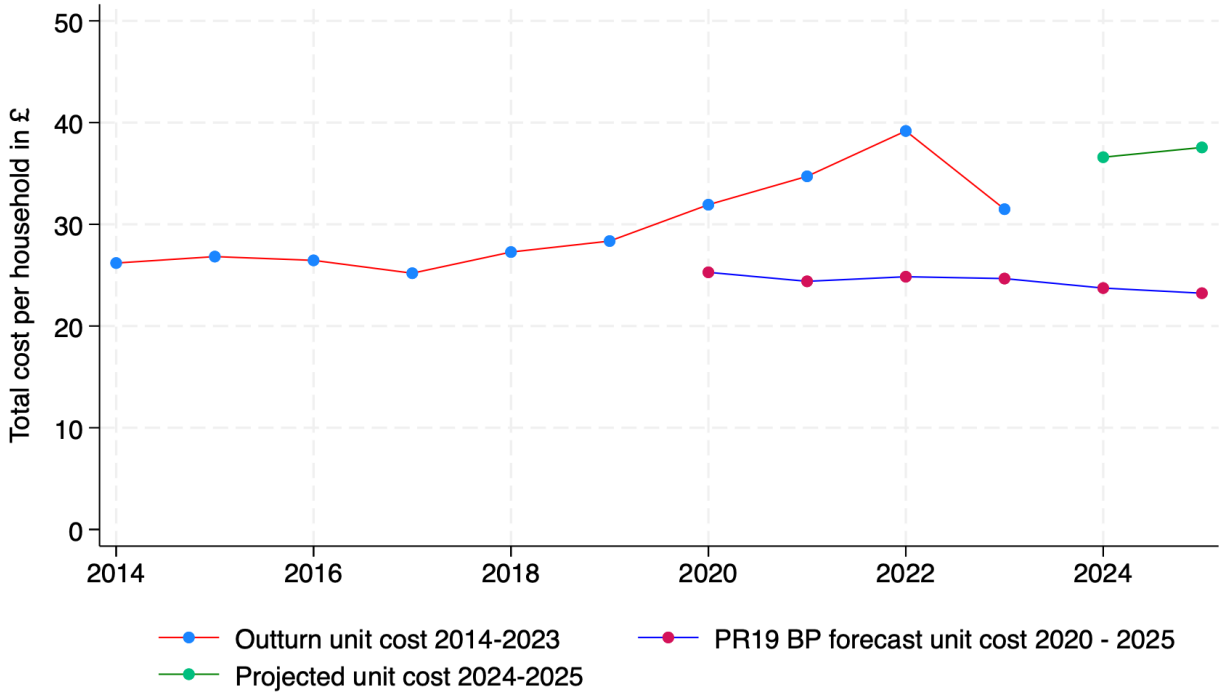


Figure 6 illustrates a similar experience for Yorkshire Water, a second example of an UQ-company. Similar to Southern Water, Yorkshire Water experienced higher than forecast costs early in AMP7, with costs failing to return to the levels forecast at PR19. The updated view of costs for the final two years of the AMP is that these will be over £10 per household higher than initially forecast at PR19.

Figure 6 Yorkshire Water AMP7 residential retail unit costs



Taken alone this may be seen as merely an updated view of costs. However, we observe that the pattern here is similar for the UQ companies, whilst for non-UQ companies the pattern is more in line with that shown for Wessex Water earlier i.e. the projected unit costs over remaining years of the AMP are not as out of line with what was forecast in the PR19 plans, indicating that initial forecasts may have been more accurate.

A1-2.2. Summary of findings

The key implications of analysing of companies' AMP7 retail residential expenditure can be summarised as follows:

- There is a clear overspend at industry-level; all but two companies overspent relative to their forecasts, and across all companies the average overspend was 14.8%.
- There is substantial variance across companies between their PR19 forecast costs and their updated projection of costs. In particular, there are differences between the set of companies whose plans at PR19 were considered to be UQ efficient, and those that were not. Across the non-UQ companies, residential retail costs are projected to exceed the PR19 forecasts by 10.2% whilst for UQ companies, the average gap is 26.0% greater than what they had been forecast. This constitutes an "extra" over-spend of 15 percentage points.