Comments on Enhancement Modelling

PR24 Draft Determination Representations – August 24

Consultation questions - Ofwat's proposed modelling for enhancement costs

11.2 Approach to model development and selection

2.1 Do you agree with our decision to use OLS to estimate our scheme level enhancement models?

While we recognise that companies' individual schemes are indeed not fully independent of each other, we would expect companies' specific effects to be less pronounced than for base costs, since a certain degree of heterogeneity is observed at the company-level between the different schemes. In this context, this approach has also the merits of being simpler and more straightforward than an approach with random effect models. Therefore, we agree with Ofwat's draft determination decision to use OLS for scheme level enhancement models. However, this question needs to be reassessed at the FD in light of updated BP data, to check whether random effect models lead to material improvements. If so, this would suggest that companies' specific effects are still present and need to be taken into account.

2.2 Do you agree with our decision to exclude outliers based on a Cook's distance threshold of 4/N?

We agree with the use of Cook's distance to determine exclusion of outliers in the context of Ofwat's current suite of enhancement models, as this measure is commonly used in regression analysis to determine highly influential data points. While there exists various alternatives to detect outliers, we do not think it would be proportionate to triangulate between different approaches in this context. Indeed, the aim is to quickly identify potential errors/inconsistencies in companies' business plan data or efficiency levels that are *far* outside the 'expected' range.

We support Ofwat's proposed approach to exclude outliers based on Cook's distance of 4/N. Although it is somewhat arbitrary to use a threshold of 4/N, this approach is relatively simple and sensible in this context of enhancement cost modelling where some reported values may be atypical or erroneous. However we note that another alternative could be to use the F-distribution method, which directly connects the threshold to a significance level, providing a formal test of whether an observation is influential. This methods considers the number of parameters in the model (with degrees of freedom $(k = 1, n - k - 1)$ and a confidence level of 50%.[1]

2.3 Do you agree with our approach to setting an efficient enhancement expenditure allowance for outlier schemes?

We agree, in principle, with Ofwat's proposed approach to determining cost allowances based on the estimated coefficients of the regression analysis (excluding these outliers) *if* the company has not provided evidence justifying why costs would be significantly higher for certain schemes. However, when companies have provided additional evidence to demonstrate that these schemes lead to

[1] See for example: Aguinis, et al., (2013). Best-Practice Recommendations for Defining, Identifying, and Handling Outliers. Organizational Research Methods. 16. 270-301. 10.1177/1094428112470848.

higher efficient costs, it is appropriate to depart from the modelled approach and rely on deep dives as proposed by Ofwat.

There may be differences in how companies and Ofwat perceive the materiality threshold for demonstrating the higher cost profile of certain schemes. However, we expect these cases to be handled individually, with consultation from the relevant companies

2.4 Do you agree with our decision to apply the PR19 log-bias adjustment to address log-bias (where relevant)?

We agree with the use of the PR19 approach. Applying the log-bias adjustment mitigates the intrinsic bias caused by using log-transformed data. This ensures that cost predictions from a log-transformed model are accurate and not systematically underestimated.

2.5 Do you agree with our decision to set the efficiency benchmark at the company level instead of scheme level?

We fully agree with this approach. Applying an efficiency challenge at a scheme level (or to individual models) would raise a significant risk of setting an unachievable target—it would establish a level of efficiency not achieved by any company.

It would also be inconsistent with Ofwat's base cost modelling where, despite a higher degree of certainty of the estimated efficiency, the catch-up efficiency benchmark is not calculated and applied individually for each model. Instead, Ofwat applies the base catch-up efficiency benchmark at an aggregate level. Aggregating first companies' predicted costs and comparing them with their business plan requests is also more suitable when determining the efficiency benchmark for enhancement cost models.

11.3 Storm overflows

3.1 Do you agree with our approach to assessing grey and grey-hybrid storage storm overflow enhancement costs?

Overall, we are supportive of the approach proposed by Ofwat in modelling enhancement costs relating to storm overflows.

First, we agree on the separation of modelling for grey and green solutions. These present clear technical differences, which limit the degree of comparability. Furthermore, we also acknowledge differences in data quality and availability, which would not support the same modelling approach for all solutions.

Second, we also agree with the decision to undertake scheme-level modelling, as it provides a significantly larger and more robust dataset compared to company-level analyses. This represents a significant improvement from the PR19 approach, especially considering the special treatment for outlier schemes which enhances further the accuracy of the modelling.

11.4 Phosphorus removal

4.1 Do you agree with our approach to assessing phosphorus removal enhancement costs?

For the same reasons set out in our response to question 3.1 above for storm overflows, we fully support Ofwat's PR24 approach to model p-removal costs at the scheme level. This represents a significant improvement over the PR19 approach at a company level with one single observation by company. The fact that the replication of the PR19 approach led to counterintuitive results is also supportive that Ofwat has adopted the right approach by moving towards a more disaggregated assessment.

The use of a larger and more robust dataset enables Ofwat to capture the specificities of the PR24 programme and its additional complexity compared to PR19.

We have great confidence in the statistical performance of the four models proposed by Ofwat and the ability of each individual cost driver to reflect the desired effect:

- population equivalent served as the key scale driver – we expect a higher population equivalent served to translate into higher costs (all other things being equal);
- enhanced phosphorus permit (with or without the squared term) – we expect tighter permit levels to translate into higher costs (all other things being equal) so it is important to capture the step change in permit levels compared to PR19;
- historical phosphorus permit – we expect higher historical permit levels to translate into higher costs (all other things being equal) as it means that solutions have not been yet developed to upgrade these sites to a new enhanced level.

We understand the rationale behind the dummy capturing a presumed step change around the technically achievable limit of 0.25 mg/L, but we do not have any specific comment on the exact threshold retained by Ofwat, as it remains somewhat arbitrary.

We also would like to emphasis on the appropriateness of a weighting given to both historical and enhanced permit levels. Conceptually, the same issue applies to leakage performance as it is cheaper and easier to reduce leakage levels from x% when a company operates in the lower quartile of the industry than it is when it operates in the upper quartile. However, Ofwat ignores this in its enhancement modelling of other leakage expenditure, considering that any company should be able to reduce leakage levels beyond its base targets based on a uniform and constant £1.1m by Ml/d improvement delivered. We consider Ofwat's approach on leakage is erroneous and inconsistent with its approach to p-removal.

Regarding Ofwat's proposed 50/50 weighting between forecast and historical models, we believe this strikes the right balance. It ensures cost efficiency is maintained for PR24 by using PR19 performance as an indicator, while also accounting for the naturally higher costs in AMP8 due to the increased treatment complexity (with tighter permit levels), and smaller schemes with fewer opportunities for economies of scale. This 50/50 weighting means that an additional catch-up efficiency challenge would be disproportionate and unjustified, as the equal weighting assigned to the PR19 data already serves as the catch-up efficiency challenge, leading to a 25% cost challenge to PR24 proposals.

4.2 Do you agree with our approach to addressing the implementation issues associated with modelling phosphorus removal enhancement costs?

We are supportive of Ofwat's reasoning for applying post-modelling adjustments.

First, we agree that excluding costs beyond 2029-30 from the modelling would likely cause distortions and reduce its accuracy due to overlap with cost drivers for AMP8 beyond. However, these costs ultimately cannot naturally be sustained within the proposed AMP8 allowance. We therefore support Ofwat's proposed reconciliation factor, which is determined by a ratio between the total requested amount for AMP8 and the costs included in the modelling. All other things being equal, this would result in a ratio lower than 1, justifying a negative adjustment on that basis.

Second, while certain specific schemes are not suitable for benchmarking and therefore justify their exclusion from the modelling, they must ultimately be re-introduced in some way into the proposed AMP8 allowance, as these costs will still be incurred during AMP8. Ofwat suggests applying the same methodology as above for costs incurred beyond AMP8 to derive the reconciliation factor for *dropped* schemes. All other things being equal, this would result in a ratio higher than 1 for companies affected by dropped schemes in the modelling, justifying a positive adjustment on that basis. We are aware of the shortcomings of this approach, as it assumes that companies are as efficient for these schemes as they are across the rest of their p-removal programme. However, we do not see a more robust alternative to address this issue. Individual assessments would likely be disproportionate and would not guarantee a more precise assessment. We believe that the average efficiency of a company across the large majority of its p-removal programme is a good proxy for determining allowances for dropped schemes.

Third, we are also supportive of the reconciliation factor for addressing minor discrepancies between companies' scheme level BP data (CWW19) and aggregate BP data (CWW3, CWW12 and CWW17). In any case this issue is likely to be fully resolved at the FD stage, following the incorporation of Ofwat's updated guidance.

11.5 Growth at sewage treatment works

5.1 Do you agree with our approach to identifying overlap with base costs so that customers do not pay for non-compliance with existing permits?

We understand that these adjustments are limited to a small proportion of the total number of schemes. When there is sufficient evidence that the requested expenditure is designed for accommodating previous non-compliance in existing sites, we agree that it appears sensible to make some form of adjustments to purely limit the assessment to requested growth costs.

5.2 Do you agree with the models we have selected to explain differences in efficient growth at STWs enhancement costs?

We fully support Ofwat's decision to exclude growth at STWs from base cost models as we think the absence of any cost drivers capturing the required step change in costs created significant limitations at PR19 on the ability of the models to properly fund the expected growth.

The proposed approach for PR24 is sensible and we welcome the triangulation between two types of models that both have merit to capture AMP8 cost requirements for growth at STWs.

The nature of these costs naturally imply a wider range of unit costs for the industry than for other cost areas, so it important to rely on key cost drivers that are able to capture genuine differences in efficiency between companies. It is also important to exclude outlier schemes that would likely bring significant noise in the econometric modelling. Overall we find that the cost drivers selected by

Ofwat provide a good proxy to capture cost differences between companies and estimate AMP8 cost requirements based on the proposed scope of work.

We also note that the cost predictions from the two models are quite similar. This is reassuring because significant differences between the models could raise concerns about their ability to accurately estimate companies' cost allowances.

5.3 Do you agree with our approach to adjusting modelled allowances to account for costs incurred outside of the 2025-30 period?

We agree with Ofwat on the need to adjust cost allowances to only consider the expenditure planned to be incurred over AMP8. However, this would need to be confirmed with the affected companies to ensure that necessary investments are not postponed until AMP9 purely based on a mechanistic cost allocation process where there remains a degree of uncertainty at this stage. For example, some AMP9 investments may in fact be required earlier than expected and ends up overlapping with AMP8. The key is to allow a sufficient degree of flexibility in the allowed cost profile.

5.4 Do you agree with our approach to adjusting allowance to account for past under-delivery?

We agree, in principle, to make an adjustment for past under delivery at PR14 and PR19. We acknowledge that our proposed DD allowance for PR24 is higher than our BP request.

Whilst there are variations between our PR14 and PR19 business plans and delivery in responding to growth, we consider that the proposed adjustment for past under-delivery provides a reasonable allowance viewed alongside our external cost benchmarks and our bottom-up cost build-up

However, conceptually, we disagree with the implementation of the proposed adjustment as it is not based on historical funded levels but rather on companies' requested figures, which is not necessarily consistent. For PR19, our business plan request for growth at sewage treatment works (STWs) was indeed £190.64m but we were only funded for £97.28m at the PR19 redeterminations, so the additional £93.36m adjustment applied on top of the £97.28m is artificial and not based on our actual PR19 allowances. To calculate the implicit allowance we have simply used Ofwat's methodology for cost adjustment claims and calculated the difference between our original PR19 modelled base cost allowances and a scenario where PR19 CMA models are re-estimated by excluding growth at STWs from the regression analysis.[2]

[2] As per example 1 in Ofwat (2022), 'Creating tomorrow, together: Our final methodology for PR24. Appendix 9 Setting expenditure allowances', December, p. 160.

11.6 Bioresources industrial emissions directive (IED)

6.1 Do you agree with our approach to setting efficient IED secondary containment, tank covering and other IED cost allowances?

At this stage we do not see a better alternative to assess these costs, so we are supportive of Ofwat's proposed approach. We agree with Ofwat that the selected cost drivers perform better in the modelling, as we find that alternative cost drivers such as sludge produced or tank volume do not improve the modelling of secondary containment or tank covering costs. Modelling other IED costs using a simple unit cost extrapolation indexed to the volume of sludge produced seems reasonable, given the difficulty of developing a specific econometric model. Unlike other areas where the degree of confidence in the estimated cost predictions are much greater, for example p-removal or storm overflows, this cost area may warrant reconsidering modelling options at the FD stage to determine whether an improved version of the modelling can be developed or if updated BP data confirms the current conclusions.

11.7 Supply interconnectors

7.1 Do you agree with our approach to assessing supply interconnector enhancement costs?

There is scope to improve Ofwat's approach to modelling supply interconnectors, given the relatively simplified model specifications used by Ofwat to assess the efficiency of these schemes.

Ofwat has used a relatively simple model, with only two cost drivers, to predict cost allowances for what are complex engineering projects. This means that when the costs of a project deviate from the predicted value, it is a matter of conjecture if this is due to inefficiency or the unique characteristics of a particular project.

Costs for a pipeline can vary due to topographic factors (e.g. road crossings, urban/suburban landscapes), or features of the pipeline itself (e.g. material, diameter, pumped vs gravity). Data on this broad range of factors is available to Ofwat.

Furthermore, among the Ofwat-included schemes, there is a wide range of costs and project sizes, not to mention complexities. Scale and complexity factors have not been included within Ofwat's model.

While there is scope to improve the modelling to better reflect scale and complexity factors, we are nonetheless supportive of Ofwat's proposed approach to incorporate, within the econometric modelling, 14 additional schemes that fall outside the supply interconnectors programme (as they relate to the Water Framework Directive, resilience or Direct Procurement for Customers interconnector schemes). We believe these additional schemes enhance the overall robustness of the econometric modelling, helping to avoid relying on just 18 observations. Although these 14 schemes are not supply-interconnectors per se, they are still informative for estimating the cost relationship with the two simple cost drivers selected by Ofwat—namely the length of the interconnector and the expected or established Ml/d benefit.

Thereafter, we also support Ofwat's exclusion of these schemes when evaluating the efficiency of the supply interconnector programme as it is more appropriate to reallocate these different interconnector schemes to their respective assessed areas. This reassignment is necessary to avoid arbitrarily impacting the various post-modelling adjustments performed by Ofwat, such as the log-bias correction or the application of the AMP8 reconciliation factor, which we fully support in our response to question 4.2 regarding p-removal. We note that this is the option retained by Ofwat in its

draft determinations, which we fully endorse. Indeed, unlike supply-interconnectors schemes, the need for other types of schemes is not fully established, as companies have a greater degree of flexibility when choosing between options. Assigning a weighting to these schemes in post-modelling procedures does not seem appropriate in this context. This is because such considerations would create an overlap between the estimated cost-efficiency and the engineering rationale of these schemes, i.e. whether the need is fully established or if alternative options would be preferable. Overall, we believe Ofwat has found the right balance by enhancing the precision of the econometric modelling while limiting the scope to supply-interconnectors schemes for post-modelling procedures in determining companies' individual cost efficiencies in this area.

Finally, we reiterate our support for the rationale of log-bias corrections, as explained in our response to question 2.4. However, we note an inconsistency in Ofwat's methodology, as the correction is applied after the AMP8 reconciliation factor instead of before. Since the purpose of log-bias corrections is to adjust cost predictions that have been transformed using a logarithmic function to reduce bias, this correction should be based on the exact modelled/requested costs for these 18 supply interconnectors schemes, not on a transformed version. We believe this may not have been intentional, as Ofwat has correctly applied the log-bias adjustment in the only other relevant area: storm overflows. For storm overflows, it is clear that the reconciliation factor is applied as the final step in the determination of cost allowances, and not before the application of the log-bias correction, which better aligns with accepted econometric standards.[3] We therefore request that Ofwat correct this inconsistency at the final determinations by applying the log-bias correction before the AMP8 reconciliation factor.

11.8 Sanitary parameters

8.1 Do you agree with our approach to assessing sanitary parameters enhancement costs?

Overall, we are supportive of the approach proposed by Ofwat in modelling sanitary parameters enhancement costs. While not ideal, it appears to be the best approach possible with the current data available. Given the existence of two different data sources, the robustness of the proposed econometric modelling will need to be confirmed at the FD to ensure that the right data has been used for all companies.

We agree with the use of population equivalent served as the sole independent variable. In particular, we find that the presence of three separate consent levels for complexity (ammonia, suspended solids and BOD) hinders the possibility of including additional drivers, analogously to what is done in the case of p-removal. While this approach oversimplifies the expected relationship, capturing the complexity at a more granular level appears to be challenging in this context.

The difficulties arise not only from the necessity of aggregating the three consent levels into one driver, so as to avoid the inclusion of six separate variables to control for the determinands' historical and enhanced consents, but also from more technical issues, including:

- The presence of markedly different content levels of sanitary determinands in untreated wastewater, which vary depending of the network and site's specific features. This leads to additional complexity when estimating the 'replacement consent' to be used when a site presents no consent.

[3] See Ofwat (2024), 'PR24-DD-WW-Storm-Overflows.xlsx', June, tab 'Allowance – Total', column M.

- The fact that not all sites perform the same spectrum of activities. For example, ammonia is generally not treated in smaller-size schemes. This further reduces the comparability of observations when controlling for the level of consent.

Additionally, we agree with the exclusion of outliers and near-zero observations, as we also find their inclusion to lead to counterintuitive results.

8.2 Do you agree with our approach to addressing the implementation issues associated with modelling sanitary parameters enhancement costs?

For the same reasons highlighted above in our response to question 4.2 regarding p-removal, we agree with Ofwat's proposed post-modelling adjustments for sanitary parameters.

11.9 Metering

9.1 Do you agree with our approach to assessing new meter installation and meter upgrade costs?

We agree with Ofwat's decision to separate metering installation and upgrades cost, given the differences in unit cost.

Regarding the modelling of new installations, we disagree with a few areas in Ofwat's proposed approach.

First, we disagree with the decision to implement a panel data model due to the lack of compelling evidence and limited engineering rationale supporting non-constant returns to scale. This is illustrated by the original regression model with an estimated coefficient very close to one, 0.98. For example, we note that Ofwat's consultation models that assessed bioresources on a total cost basis have now been completely removed despite a scale driver coefficient as high as 1.13 (and statistically significant at the 1% level). This suggests that a unit cost model may be more suitable for new installations, which would also better align with the engineering rationale.

This could take the form of a simple unit cost modelling based on the median ratio excluding outliers such as SEW, HDD or SES.

Alternatively, we believe there is also merit in using a single pooled econometric unit cost model that includes population density and meter penetration to account for additional factors besides scale. We consider that incorporating population density and total meter penetration is relevant, given the supporting engineering rationale for their inclusion. These options were also explored by Ofwat but found to be inconclusive due to the use of a panel data modelling form. We observe large variability across the industry on these metrics and our preliminary findings show that, when modelled on a unit cost basis with one observation by company, these variables are statistically significant and positive as per economic intuition.

For new installations, we therefore recommend Ofwat to use one of the following approaches at the FD:

- a simple unit cost model with the use of a median ratio;

- a pooled econometric unit cost model with density and meter penetration as cost drivers, given their strong engineering rationale;
- a triangulated approach.

To avoid favouring either of the first two options, we recommend using a triangulated approach.

Separately, we note that companies like SEW, SSC, HDD or SES seem to incur atypical low/high costs in this area. Their suitability for benchmarking under Ofwat's current proposed modelling is unclear and will need to be reconsidered at the FD stage if this remains the case. We believe this variation is likely due to differences in the types of meter installations that will be undertaken during AMP8. We provided this information to Ofwat in response to query OFW-OBQ-ANH-055, where we detailed the unit costs for each type of meter installation (internal, external requiring digs, and external not requiring digs). The highest unit cost for external meters requiring digs was four times greater than the lowest unit cost for external meters not requiring digs.[4] Given that Ofwat benefits from disaggregated data at the industry-level, we request that Ofwat evaluate whether a more detailed assessment by type of meter installation could improve the robustness of the modelling.

Regarding meter upgrades, we note that the data constructed by Ofwat in its Stata .do file leads to constant unit rates across *all* years for *all* companies. It is, therefore, clear that a panel data model is not appropriate, as time has no effect.

Therefore, we recommend Ofwat to use a simple unit cost model based on the median ratio, for the similar reasons expressed for new installations.

We have also assessed a pooled cross-sectional unit cost model as for new installations with density and meter penetration as cost drivers, but results were inconclusive. So a unit-cost modelling appears to be the most appropriate option to reflect the constant unit costs assumed by the entire industry over AMP8.

Similar to the point raised above regarding new meter installations, we believe that a significant proportion of industry-wide variations is due to differences in how companies allocate infrastructure costs — whether as data provision services (OPEX) or network building (CAPEX). Therefore, we consider a separate assessment at the FD to be justified for companies that have adopted different approaches.

9.2 Do you agree with our decision to assess smart infrastructure costs within the meter installation and meter upgrades models?

We support the inclusion of smart infrastructure costs within the meter installation and meter upgrades modelling. This helps mitigating the variation arising from potential inconsistencies across companies in allocating these costs. However, we note that for the meter upgrades model, smart meter infrastructure is the only source of variation across time, as the proposed unit costs are otherwise constant at the company-level. Therefore, this must be noted if evaluating a model on a unit cost basis in a cross-sectional sample instead of panel data.

[4] 639.58 vs 157.41.

11.10 Lead

10.1 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 the number of pipes replaced is the key driver.

We also agree with Ofwat that the average length of pipes would add little additional explanatory power to the model, as it is relatively uniform across companies.

We are however not fully aligned with Ofwat's position on density. Although the current dataset used for the modelling does not clearly establish such relationship, from an operational perspective we would expect a U-shaped relationship between the sparsity (or density) of pipe locations and the cost of replacing or relining them. That is, workloads in both remote and very dense, urban areas are expected to be relatively more costly.

10.2 Do you agree with our approach to triangulating between the median unit cost and an econometric model?

No, we disagree with the approach. We believe *all* the weight should be placed on a more transparent and robust econometric modelling approach. That is, it is not intuitive for Ofwat to assume that there are no economies of scale, when the modelling empirically shows that to be the case.

We also do not agree with the functional form of Ofwat's model. Given that Ofwat is effectively using a unit cost model (cost predicted by volume of activity), the univariate panel structure adds no value.

Instead, the year-on-year panel in all likelihood only introduces noise. The within company variation in unit costs seemingly observed over time is likely due to mismatches between when these companies expect to start making the relevant expenditure (e.g., in year t) relative to when the corresponding volumes are delivered (e.g., year t+1). We do not believe that companies are in fact expecting their unit costs to vary significantly over time, e.g. due to changing nature of the workloads. Practically, we also do not expect our unit costs to vary between £1,526.49 per pipe in 2026/27 and reduce to £1,057.86 in 2029/30[5]—Ofwat's modelling approach simply misinterprets the slightly different time profiles of when we start making the relevant expenditure relative to when we deliver and finalise the eventual delivery.

A simple collapsed model, regressing companies total costs over total number of pipes replaced, would thus be more appropriate (unless Ofwat believes that there are other time-varying factors that further need to be included as cost drivers).

[5] As suggested by the '*CP_Unit_costs analysis'* sheet in Ofwat's lead enhancement model, *PR24-DD-W-Lead.*