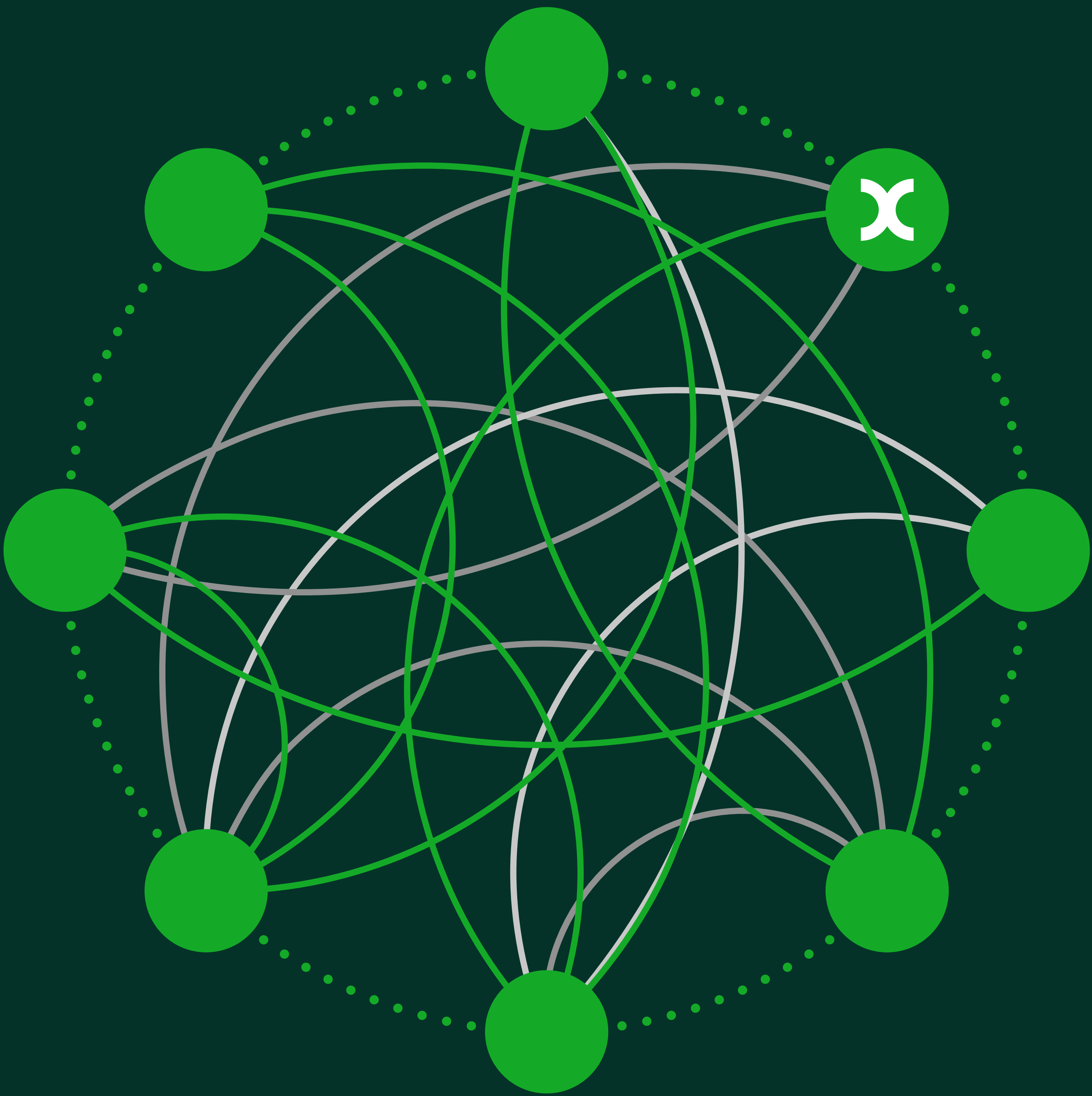
Base cost modelling at PR24

—

Prepared for Yorkshire Water Services

27 August 2024



Contents

Oxera Consulting LLP is a limited liability partnership registered in England no. OC392464, registered office: Park Central, 40/41 Park End Street, Oxford OX1 1JD, UK with an additional office in London located at 200 Aldersgate, 14th Floor, London EC1A 4HD, UK; in Belgium, no. 0651 990 151, branch office: Spectrum, Boulevard Bischoffsheim 12–21, 1000 Brussels, Belgium; and in Italy, REA no. RM - 1530473, branch office: Rome located at Via delle Quattro Fontane 15, 00184 Rome, Italy with an additional office in Milan located at Piazzale Biancamano, 8 20121 Milan, Italy. Oxera Consulting (France) LLP, a French branch, registered in Nanterre RCS no. 844 900 407 00025, registered office: 60 Avenue Charles de Gaulle, CS 60016, 92573 Neuilly-sur-Seine, France with an additional office located at 25 Rue du 4 Septembre, 75002 Paris, France. Oxera Consulting (Netherlands) LLP, a Dutch branch, registered in Amsterdam, KvK no. 72446218, registered office: Strawinskylaan 3051, 1077 ZX Amsterdam, The Netherlands. Oxera Consulting GmbH is registered in Germany, no. HRB 148781 B (Local Court of Charlottenburg), registered office: Rahel-Hirsch-Straße 10, Berlin 10557, Germany, with an additional office in Hamburg located at Alter Wall 32, Hamburg 20457, Germany.

Base cost modelling at PR24

Introduction

Yorkshire Water (YWS) has commissioned Oxera to review the econometric models that Ofwat used at the PR24 draft determination (hereafter, the 'DD models') to assess base expenditure requirements. The DD models are broadly similar to those that Ofwat presented at the PR24 base cost modelling consultation in April 2023,[1] although Ofwat has taken on board some of the recommendations when refining those models for the draft determination (DD). When using these models to assess base expenditure requirements, Ofwat has applied an upper-quartile (UQ) benchmark across all price controls, which is aligned with the benchmark applied by the Competition and Markets Authority (CMA) in the PR19 redetermination. Ofwat notes that it will consider the stringency of the benchmark at the final determination (FD), following the possible inclusion of 2023/24 data and its consideration of stakeholder responses to the DD.[2]

In this report, building on our previous submissions,[3] we propose targeted improvements to the DD models that can better capture industry-wide cost pressures. These improvements are based on a combination of empirical evidence, regulatory precedent and additional operational arguments provided by YWS. As part of our review of the DD models, we have also undertaken an investigation of the reliability and uncertainty associated with the models and data using objective, scientific methods. Following the principles outlined by the CMA in the PR19 redetermination—namely that the benchmark should be informed by the quality of the econometric models—we have used the scientific methods that we outlined in our report alongside YWS's business plan submission to assess Ofwat's choice of benchmark.[4]

Our review of the DD models, our proposed improvements and the results from these for YWS on wholesale water (WW), wholesale wastewater network plus (WWNP), bioresources (BR) and residential retail are outlined in the sections below. The models estimated in this report are based on a combination of Ofwat's base modelling datasets (which includes data for 2012–23 for the wholesale models and 2014–23

[1] Ofwat (2023), 'Econometric base cost models for PR24', April.
[2] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances', July, p. 24.
[3] See Oxera (2023), 'An assessment of Yorkshire Water Services' base cost requirements', September.
[4] The methodology is outlined in more detail in Appendix A1.

for the retail models) and the latest data derived from companies' Annual Performance Reports (APRs), which extends the dataset by an additional year (i.e. to include 2023/24).[5]

Wholesale water modelling

Ofwat has not made any changes to its DD models in WW relative to its PR24 base cost modelling consultation. As such, the incremental improvements to the modelling that we proposed in our submission alongside YWS's business plan remain relevant for the DD. Our proposed improvements to the DD models are outlined in the table below.

[5] More details on how the dataset has been constructed can be found in Appendix A1.

Area	Proposed approach
Scale (TWD models)	Ofwat continues to use length of mains as the sole cost driver in its treated water distribution (TWD) models. Ofwat's primary argument against the inclusion of connected properties is that it should (in Ofwat's view) have an immaterial impact on allowances. However, **this is an incorrect statement**: YWS's allowance is c. £8m higher in TWD models that control for connected properties, and individual companies' allowances vary by -2.4% to +1.6%. Connected properties is a relevant scale driver in TWD, given that it can better capture the costs associated with population growth (e.g. network reinforcement), as evidenced by a stronger correlation between network reinforcement costs and connected properties relative to length of mains. Furthermore, since the two drivers perform similarly at a statistical level, **we control for connected properties as the scale variable in half of our TWD models.**
Treatment complexity	Ofwat continues to adopt the natural logarithm of weighted average treatment complexity (WAC) **despite it being a proportion variable like the other treatment complexity measure** the 'proportion of water treated at complexity band three and above', which Ofwat models in levels. The WAC is a **weighted proportion measure** that is unitless, while the other complexity measure is an unweighted proportion measure (also unitless).[6] To derive WAC, Ofwat assigns weight between 1 and 7 to represent the different complexity levels (hence, companies' WAC value is *bound* between 1 and 7). However, in a regression context, where scaling of data does not affect the outputs, this is equivalent to assigning weights between 0.14 (1/7) and 1 (7/7) for these complexity levels, in which case the WAC is *bound* between 0 and 1 for companies, as in a proportion variable similar to the other complexity measure. **It should be clear that the WAC is not a ratio measure (e.g. connected properties over mains length) but is a proportion variable**. Hence, the WAC and not the logarithm of it measures the average level of complexity of water. When the WAC is modelling in levels, it has a clear operational interpretation: moving 1% of water from complexity band 'x' to complexity band 'y' is associated with an increase in predicted costs of (y-x) multiplied by the coefficient. This relationship can be validated against operational expectations. When modelled in logarithm, moving 1% of water from complexity band 'x' to complexity band 'y' has no clear operational interpretation, other than in improbable cases.[7] Not only is modelling **WAC in levels more operationally intuitive and consistent with Ofwat's approach to modelling proportion variables**, it also leads to a clear **improvement in the statistical quality** of the models. **We control for WAC in levels in half of our WRP and WW models.**

[6] This compares to cost drivers that represent a ratio (e.g. properties per length of mains) or an average (e.g. weighted average treatment plant size), where the cost driver would have units. For example, the unit for properties per length of mains would be the number of properties per km and the unit for weighted average treatment plant size would load (in kg BOD5 per day).

[7] For example, if a company that currently treats all of its water at simple works (i.e. WAC=1) moves all of its water to complexity band 1 (i.e. WAC=2, a doubling of the WAC measure), this is associated with the same cost increase as a company that currently treats all of its water at complexity band 2 (i.e. WAC=3) to complexity band 5 (i.e. WAC=6, a doubling of the WAC measure). While this extreme case could potentially be validated through operational insight, the interpretation of the coefficient becomes more challenging when companies treat water at a mixture of complexity levels. This extreme case is inconsistent with Ofwat's operational expectations that there is a step-change in cost requirements at complexity band 3.

Area	Proposed approach
Topography	As during the submission phase of companies' business plans, YWS has reiterated the strong rationale for using the number of booster pumping stations per length of mains (BPSL) as a direct proxy for network topography in modelling treated water distribution and wholesale water costs. Ofwat considers that both APH and BPSL have pros and cons and has proposed to triangulate results from the corresponding TWD and WW models that include either measure. **On Ofwat's dataset, with and without the 2023/24 data, BPSL and APH appear uncorrelated. This could be because, as Ofwat notes, neither measure is necessarily a perfect proxy**. Given that neither measure is perfect, and their operational relevance has been supported by companies in the industry, they could capture different aspects of costs. In such a case, a direct test of omitted variable bias (OVB) is to assess the statistical significance and interpretability of the relevant measure directly in the model. Therefore, we have re-assessed the empirical evidence of their joint inclusion in the same model. Given the strong modelling performance of this approach, **we include APH and BPSL in all TWD and WW models**. This is a potential option for Ofwat to consider in the FD as it could help address any OVB issue in considering them separately (**as confirmed by the direct test**) that may not be offset through triangulation.

Source: * Ofwat (2024), 'PR24 draft determinations: Expenditure allowances – Base cost modelling decision appendix', July, p. 17.

Overall, our proposed changes lead to an improvement in model quality when compared with the DD models. The model fit generally improves and some coefficients that were statistically insignificant in Ofwat's models become statistically significant in ours. Where coefficients are statistically insignificant, they are close to the 10% threshold. Given that these models represent an improvement on Ofwat's DD models, they should be considered when assessing YWS's base expenditure allowances at the FD.

The predictions from Ofwat's DD models are associated with a higher degree of uncertainty than those at the PR19 redeterminations when the CMA selected the UQ as its benchmark. Our estimates, using pooled and panel stochastic frontier analysis (SFA), indicate that the UQ seems to correspond to the most challenging benchmark choice that Ofwat could reasonably assume at the FD stage. Despite the targeted improvements to the cost driver specification in our proposed models, the underlying level of uncertainty do not support a benchmark more stringent than the UQ.

The following table summarises YWS's efficient cost allowances for WW across Ofwat's DD and our proposed models.

YWS's efficient cost predictions—WW

	Submitted costs	Efficient modelled cost
PR24 DD models with APR update	2,334	1,657
Our proposed models with APR update	2,334	1,691

Note: The estimated values above include a UQ benchmark and FS/RPE assumptions as per Ofwat's DD decision, and are in £m 2022/23 prices.
Source: Oxera analysis.

YWS's efficient cost prediction under the proposed models is c. £1,691m, which is c. £34m higher than that predicted by Ofwat's DD models. That is, using better-specified models increases YWS's estimated efficient cost prediction.[8]

Wastewater network plus

Ofwat has removed ten of its 17 PR24 consultation models and included three models for sewage collection (SWC), two models for sewage treatment (SWT) and two models for WWNP. It has included the following cost drivers: (i) scale; (ii) economies of scale at sewage treatment works (STWs); (iii) treatment complexity; (iv) network topography; (v) population density; and (vi) urban rainfall. While the DD models represent an improvement on the PR24 consultation models—for example, urban rainfall is now accounted for in all relevant models—incremental improvements to the models are needed for the FD.

One key issue with Ofwat's models is that it continues to adopt the percentage of load treated in size bands 1–3 as an economies of scale cost driver. Unlike the weighted average treatment size (WATS) variable, this relies on the arbitrary threshold of 1–3 which assumes that all STWs in these size bands have the same level of efficient unit costs, and that all STWs in size bands 4 and above have the same level of efficient unit costs. Moreover, the WATS driver outperforms the 'percentage of load treated in size bands 1–3' driver from a statistical perspective. Therefore, we consider that models that control for the percentage of load treated in size bands 1–3 should be removed from the modelling suite (and not given the same weight as models with WATS).

[8] While a material efficiency gap still exists under our proposed models, this efficiency gap is reduced significantly once the post-modelling adjustments (e.g. relating to mains replacement) are factored in. See Oxera (2024), 'Cost adjustment claims', August.

We note that Ofwat's models do not account for several operationally relevant drivers of expenditure, particularly phosphorus removal (P-removal) activity and network characteristics (e.g. combined sewers). Ofwat accounts for increased P-removal activity through a post-modelling adjustment and we consider that a post-modelling adjustment is likely to be appropriate for this activity, given that it is comparatively 'new' and is therefore difficult to model on outturn data. We consider that network characteristics should be captured directly within the econometric models and, if network characteristics are not captured in the models, companies should receive necessary funding through other means (e.g. cost adjustment claims, CAC). Our review of Ofwat's post-modelling adjustment for P-removal and Ofwat's rejection of YWS's CAC relating to combined sewers is covered in a separate report.[9]

The uncertainty associated with the predicted costs under Ofwat's DD models for SWC and SWT are broadly comparable to the equivalent models at the PR19 redeterminations where the CMA applied a UQ benchmark. The confidence intervals for the WWNP models suggest these estimate companies' costs with a lower degree of uncertainty than SWC and SWT. The pooled SFA analysis shows that the level of inefficiency generally indicates that a UQ benchmark may be appropriate, whereas the panel SFA models indicate that there is no statistically significant inefficiency in the sample (which suggests that a less stringent benchmark would be more appropriate). While the level of uncertainty improves under our specifications in most cases, the scientific methods considered do not support a benchmark more stringent than the UQ.

Using the APR 2024 data share provided by YWS, we have updated Ofwat's DD cost assessment dataset, and the subsequent cost driver forecasts. The following table summarises YWS's efficient cost allowances for WWNP across Ofwat's DD and our proposed models.

YWS's efficient cost predictions—WWNP

	Submitted costs	Efficient modelled cost
PR24 DD models with APR update	1,723	1,752
Our proposed models with APR update	1,723	1,752

[9] Oxera (2024), 'Cost adjustment claims', August.

Note: The estimated values above include a UQ benchmark and FS/RPE assumptions as per Ofwat's DD decision, and are in £m 2022/23 prices. The values prior to the APR update also do not incorporate any changes made to the cost assessment dataset or, in the case of the proposed models, reflect the final proposals.
Source: Oxera analysis.

YWS's efficient cost prediction under the proposed model is c. £1,752.1m, which is c. £0.6m higher than that predicted by Ofwat's DD models.

Bioresources

Ofwat has only proposed unit cost models for BR, arguing that this is more consistent with how the average revenue control is designed and will help to support the bioresources market. This represents a departure from the PR24 modelling consultation, where six of the ten proposed BR models were modelled on a total cost basis. In addition to scale (which is captured through the unit cost modelling), Ofwat also controls for STW-level economies of scale.

The BR models are particularly parsimonious when compared with Ofwat's other wholesale and retail models, which may partly explain why the model fit is particularly low. Nonetheless, the cost drivers are statistically significant across the models, and our investigation has not found model specifications that are clearly superior to Ofwat's DD models.

Our benchmark analysis shows that BR costs are estimated with a high degree of uncertainty under the DD models, and the SFA modelling suggests that the models cannot detect any statistically significant inefficiency. Given the parsimony of the models and the large uncertainty associated with companies' cost predictions, a UQ benchmark is likely to be overly challenging.

The following table summarises YWS's efficient cost allowances and the cost gaps with its submitted costs for BR.

YWS's efficient cost predictions—BR

	Submitted costs	Efficient modelled cost
PR24 DD models with APR update	338	359

Note: The estimated values above include a UQ benchmark and FS/RPE assumptions as per Ofwat's DD decision, and are in £m 2022/23 prices.

Source: Oxera analysis.

YWS performs well under Ofwat's bioresources models, with an efficient cost allowance that is higher than its submitted costs by c. £21m after the APR update.

Residential retail

Ofwat's DD models control for the following characteristics:

- economies of scale, through the inclusion of connected properties in one of its other cost (ROC) and three of its total cost (RTC) models;
- revenue at risk, through average bill size in all of its bad debt (RDC) and RTC models;
- propensity to default, through measures of deprivation (percentage of households with payment default, income deprivation score) in its RDC and RTC models;
- type of customer, through the proportion of dual-service households in one of its ROC models;
- the effect of the COVID-19 pandemic, through time dummies for the years 2020 and 2021 in all of its RDC and RTC models.

Compared with the PR24 base cost modelling consultations, Ofwat excluded the average number of county court judgments because this measure performed poorly on the latest dataset (the coefficient was weakly statistically significant compared with the other measures of deprivation). Consequently, the models that included this variable were removed, reducing the total number of models from 11 to eight.[10]

Once the modelling period is extended to include the latest APR data, all cost drivers remain statistically significant and directionally aligned with operational expectations.[11]

COVID-19 dummies are not needed and should be dropped

The primary limitation with the DD models is that Ofwat has continued to include COVID-19 dummies in all of its RDC and RTC models. Ofwat

[10] The eight remaining models were (i) two bad debt cost models (RDC), (ii) two other cost models (ROC), and (iii) four total retail cost models (RTC).
[11] A more detailed summary of the models and their performance is available in Appendix A5.

notes that the inclusion of these dummies is required to ensure the optimal statistical performance of its models. Ofwat has stated that it would reconsider the inclusion of COVID-19 dummies when incorporating the additional outturn data.

We previously noted that the time dummies are blunt instruments as they capture myriad effects and not just the specific increase in doubtful debts in those years.

On the augmented dataset, we have explored targeted approaches to account for the COVID-19 year anomalies that do not require strong and untested assumptions. From our review, we consider three options that are superior to Ofwat's time dummies:

- smoothing doubtful debt costs over the modelling period, in line with Ofwat's treatment of depreciation;
- using linear interpolation to impute the doubtful debt costs in 2020 and 2021;
- replacing the doubtful debt costs in 2020 and 2021 using the average doubtful debt costs in 2017–19 and 2022–24 (i.e. data outside of the COVID-19 years).

In addition to avoiding strong and untested forward-looking assumptions that are embedded in time dummies, these models perform well against Ofwat's modelling criteria—the estimated coefficients are statistically significant and retain the expected sign, and the model fit is comparatively high.

The first option of smoothing doubtful debt costs over the entire modelling period is in line with the approach we proposed alongside YWS's business plan submission. When this method is applied on the augmented dataset, the model fit is significantly better than in Ofwat's models.[12] The other two approaches outlined above are viable options that offer more consistent outcomes compared to Ofwat's PR24 DD models.

Moreover, on the augmented dataset, as Ofwat had envisioned in the DD, we consider that the COVID-19 dummies could potentially be removed from the model specifications without any special consideration of doubtful debt costs. Doing so leads to coefficients that

[12] The adjusted R-square is higher on average by c. 10 percentage points in RDC models and c. 3 percentage points in RTC models. It is important to interpret these comparisons with caution, given the different nature of the dependent variables.

remain directionally intuitive and (in most cases) statistically significant.[13] Note that the model fit reduces by construction, but the deterioration in model fit is not material.

The only area where there may be a perceived improvement in model quality from incorporating COVID-19 dummies is in relation to the RESET. Specifically, Ofwat's DD models 'pass' the RESET at standard thresholds, while models that exclude the COVID-19 dummies typically fail the RESET. However, Ofwat's models generally fail the RESET when the models are estimated over the pre-COVID-19 years, i.e. 2014–19 (the time period used at PR19) or the post-COVID-19 years, i.e. 2022–24. That is, Ofwat's models that fail the RESET test is not an issue related to the inclusion or exclusion of COVID-19 dummies; rather, it is a general issue with Ofwat's model specification, which suggest that the COVID-19 dummies could be capturing isolated data anomalies instead.

Upon exploring further, we noticed that when data from two companies (United Utilities and Dŵr Cymru) are excluded from the models, the models without COVID-19 dummies pass the RESET.[14] A reasonable interpretation of this finding is that **the inclusion of COVID-19 dummies, rather than effectively controlling for the spike in doubtful debts during the COVID-19 years (i.e. structural break), is instead capturing isolated outliers (that the RESET mistakes for non-linearities)**. In this context, the COVID-19 dummies are not necessary for improving model quality and may inadvertently capture a range of unintended diverse effects, while artificially depressing forward-looking allowances (as the outturn-based models are extrapolated to derive AMP8 allowances).

Finally, companies' cost predictions using the PR24 DD models are estimated with a higher degree of uncertainty than the equivalent models at PR19. Moreover, both the pooled and panel SFA are unable to capture any statistically significant inefficiency in any of the bottom-up models, supporting the decision to place more weight on the top-down models.[15]

The following table presents the allowances (both in 2022–23 and nominal prices) obtained from: (i) PR24 DD models; (ii) our proposed

[13] The income deprivation score variable loses statistical significance, but its directional sign remains aligned with operational expectations.
[14] A more detailed summary can be found in Appendix A5.
[15] Our proposed models decrease the uncertainty around the estimations for bad debt costs but the corresponding SFA models are still unable to capture any statistically significant inefficiency.

models (i.e. smoothing doubtful debt over the modelling period); (iii) PR24 DD models excluding COVID-19 dummies.

YWS's efficient cost predictions—residential retail

	Submitted costs	Efficient modelled costs (2022/23 prices)	Efficient modelled costs (nominal)
PR24 DD models with APR update	446.0	437.9	494.1
Our proposed models with APR update	446.0	459.7	518.7
PR24 DD models with APR update and without COVID-19 dummies	446.0	462.1	521.5

Notes: Figures are estimated using a UQ benchmark, with a 1% ongoing efficiency and 0.23% RPE.
Source: Oxera analysis.

As presented in the table above, our proposed model provides a material improvement of c. £21.8m (in 2022/23 prices) compared with Ofwat's PR24 DD models.[16] The same holds true for Ofwat's PR24 DD models without COVID-19 dummies, showing a material improvement of c. £24.2m (in 2022/23 prices).

Summary

Our proposed models across price controls result in a more robust assessment of YWS's efficient expenditure requirements, relative to the PR24 DD models.

The following table summarises YWS's efficient cost predictions across the different price controls.

[16] The two other proposed models also predict efficient modelled costs in the range of c. £460m (2022/23 prices).

YWS's efficient cost prediction summary

	Submitted costs	Proposed models	Ofwat's DD models
WW	2,334	1,691	1,657
WWNP	1,723	1,752	1,752
Bioresources	338	359*	359
Residential retail	446	460	438
Total	**4,841**	**4,262**	**4,206**

Note: The estimated values above include a UQ benchmark and FS/RPE assumptions as per Ofwat's DD decision, and are in £m 2022/23 prices. * We have not made any proposed changes to the bioresources models, therefore we assume the efficient expenditure assessment from Ofwat's DD models to be our proposed level of expenditure.
Source: Oxera analysis.

Under Ofwat's DD models, YWS's estimated cost gap is c. £635m. However, under our proposed models, this cost gap narrows to £579m. Moreover, this remaining cost gap can be robustly explained by our proposed improvements to Ofwat's post-modelling adjustments, which would result in an uplift to YWS's efficient cost prediction of c. £737m.[17] That is, YWS's business plan is assessed to be efficient under Ofwat's current assumptions on the benchmark, ongoing efficiency and RPEs.

Finally, we note that Ofwat expects companies to achieve stretching performance commitments through base expenditure allowances. However, Ofwat has not presented any detailed analysis to suggest that its performance targets are achievable through base expenditure, and our analysis of what is implicitly funded through base expenditure suggests that Ofwat's targets are overly stretching and may be undeliverable. Therefore, in combination with the proposed improvements to cost modelling (outlined in this report) and post-modelling adjustments,[18] Ofwat could consider: (i) relaxing some of the performance commitment levels; (ii) allowing additional expenditure to achieve stretching service targets; (iii) some combination of the two.[19]

[17] For a summary of our assessment of the post-modelling base cost adjustments, see Oxera (2024), 'Cost adjustment claims', August.
[18] See Oxera (2024), 'Cost adjustment claims', August.
[19] Oxera (2024), 'Addressing the disconnect between cost and outcomes', August.

A1 Methodology

A1.1 Dataset

Our modelling dataset is derived from two sources. The majority of the data relating to costs and cost drivers is derived from Ofwat's base cost modelling dataset published as part of the PR24 draft determination in July 2024.[20] This includes outturn data for all companies between 2012 and 2023 in the wholesale modelling, and outturn data for all companies between 2014 and 2023 in the residential retail modelling. The datasets also include submitted business plan data for all companies between 2024 and 2030.

We have supplemented Ofwat's cost assessment dataset with outturn data from 2024, derived from an industry data share of the 2024 APRs, provided to Oxera by YWS. This allows for the modelling period to be extended by a year from 2012–2023 (2014–2023 in residential retail) to 2012–2024 (2014–2024 in residential retail).

We have identified a potential inconsistency in the reporting of Portsmouth Water's (PRT) booster pumping stations (BN11390) in Ofwat's cost assessment dataset and its APR. That is, Ofwat has reported a value of 22 for PRT's boosters in 2022/23, despite PRT's 2022/23 APR data table stating a value of 40, which is consistent with PRT's values in previous years. [21] Furthermore, the 2024 APR data share reports a value of 23 for PRT's boosters in 2023/24.[22] Since this driver is usually stable over time, and is expected to increase according to the forecasts in PRT's business plan, we consider that it is unlikely that PRT has decommissioned almost half of its boosters, before recommissioning them in AMP8.[23] Such a difference can have an impact on the model coefficients, in particular the magnitude and significance of the boosters per length driver. This can subsequently affect the benchmarking and estimation of efficient cost predictions. Therefore, we have assumed PRT's booster pumping stations to be 40 for the

[20] See Ofwat (2024), 'Base costs – water model 1', July, https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-Base-costs-water-model-1.xlsx; Ofwat (2024), 'Base costs – wastewater model 1', July, https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-Base-costs-wastewater-model-1.xlsx; Ofwat (2024), 'Base costs – residential retail model 1', July, https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-Base-costs-residential-retail-1.xlsx.

[21] Portsmouth Water (2023), 'Annual Performance Report Data Tables 2023', tab 6B, https://www.portsmouthwater.co.uk/wp-content/uploads/2023/11/2022-23-annual-performance-report-excl-tables-3A-3Iv1.1-PRT-23Nov.xlsx, accessed 20 August 2024.

[22] Industry APR data share provided by YWS, tab 6B.

[23] Portsmouth Water (2024), 'PRT51: PR24 – Business Plan Tables – 26 January 2024 Update', tab CW5, https://www.portsmouthwater.co.uk/downloads/pr24/PRT51-Jan24update.xlsb, accessed 20 August 2024.

financial years 2022/23 and 2023/24 to ensure consistency with PRT's historical data, as well as its submitted BP forecasts across AMP8.

Some external data used in Ofwat's base cost modelling dataset is not reported in companies' APRs. For these variables, we have made the following assumptions to estimate the 2024 value.

- Across water and wastewater, weighted average density—MSOA to LAD (BN4013 & BN4015) and MSOA (BN4000 & BN4006) is based on the outturn growth rate of properties per length of mains/per sewer length in 2024 applied to the weighted average density in 2023, as per Ofwat's forecast approach.
- WATS (STWDP160) is set to the average of the last three years, as per Ofwat's approach. WATS is relatively stable over time, so we consider it appropriate to use the average of the three most recent years.
- Urban MSOA rainfall (BN4507B) is set to Ofwat's forecast for 2023/24 due to a lack of information regarding outturn rainfall data.

We have made no changes to the remaining cost driver forecasts since the majority are based on company forecasts.

A1.2 Historical mergers

During the modelling periods (2012–24 in wholesale, 2014–24 in retail), there have been three potential changes in the structure of the industry that require consideration when undertaking modelling.

First, there was a merger between South West Water (SWT) and Bournemouth Water (BWH) in 2016. Until this point, the two companies reported data on costs and outputs separately. Thereafter, costs and outputs have been reported under a single entity (SWB). In line with the approach taken by Ofwat at the PR24 draft determination, we merge the data for SWW and BWH into a single entity (SWB) in the years prior to the merger.[24]

Second, there was a merger between Severn Trent Water (SVT) and Dee Valley Water (DVW) in 2018. The merger involved the creation of two new entities: Severn Trent England (SVE) and Hafren Dyfrdwy (HDD). SVE undertakes the water, wastewater and retail services previously

[24] This approach assumes that SWT and BWH were not operationally independent prior to the merger. While this is a simplifying assumption, it may have a disproportionate impact on some companies.

undertaken in the English regions of SVT and DVW, while HDD undertakes the water, wastewater and retail services previously undertaken in the Welsh regions of SVT and DVW. As DVW was a water-only company (WOC), this merger involved the creation of a new water and sewerage company (WaSC) that is materially smaller than those in the rest of the industry.

In line with precedent from the PR19 redetermination and Ofwat's PR24 draft determination analysis, we: (i) have treated HDD and SVE as new independent companies in the WW and residential retail modelling; (ii) have combined the cost and output data for HDD and SVE into a new entity (SVH) in the wholesale wastewater modelling.

In 2023, SWB and Bristol Water (BRL) merged. As part of the CMA's decision to accept the merger, SWB and BRL are required to continue to report data on costs and outputs separately (among other aspects). In the analysis presented in this report, we treat SWB and BRL as separate and independent entities, as per Ofwat's draft determination analysis, given that the merger affects only two years of data. However, going forward, it may be appropriate to merge the data for SWB and BRL (in line with the treatment of the SWT–BWH merger), given that the two entities are no longer independent.

A1.3 Modelled expenditure

In our assessment of YWS's efficient BOTEX requirements, we have excluded cost items from the modelled cost base that are either outside management control or could provide perverse incentives with respect to cost reduction; this is consistent with Ofwat's approach at the DD. In the wholesale cost models, excluded costs include business rates; costs associated with the Traffic Management Act; costs associated with statutory water softening; abstraction charges and discharge consents; diversions (NRSWA and other non-S185 diversions); and the developer services base cost adjustment.

Alongside the base expenditure, the modelled costs also include network reinforcement expenditure and certain enhancement activities, in line with Ofwat's modelled cost definitions.

A1.4 Modelling approach

To assess YWS's efficient base expenditure, we employ econometric cost benchmarking to the historical cost assessment dataset described in the section above, as per Ofwat's DD approach. These models are then applied to the cost driver forecasts to estimate a company's cost allowance at the 'average efficiency' level. While we use Ofwat's econometric cost models at the DD, we have also proposed our own version of the models which improve upon Ofwat's models from an operational[25] and/or statistical perspective (these changes are discussed in detail in their respective sections).

We also apply the same upper-quartile (UQ) benchmark for the catch-up challenge as Ofwat in its DD (see section below on the methodology of assessing a reasonable benchmark), as well as the same frontier shift (1.0%) and RPEs to assess the final efficient cost prediction of YWS across the various controls. Since the 2024 APR update is used to extend the modelling period, the effects of the net frontier shift begin from 2024/25 (instead of 2023/24 in the DD).

A1.5 Benchmark modelling methodology

We employ two techniques to examine the precision of Ofwat's econometric modelling: confidence intervals associated with the cost estimations of the companies, and SFA (pooled and panel) as alternative benchmarking modelling.

A1.5.1 Confidence interval methodology

In the context of cost assessment modelling, we measure the precision of an econometric model through the confidence interval around its predicted values. A 95% CI indicates that there is a 95% probability that the true cost value lies within the specified range. The narrower the confidence interval the more precise is the model. This interval provides an estimate of uncertainty around the predicted costs of the companies, helping to capture the precision of the model. For instance, these play a key role in Ofwat's (and the CMA's) merger impact assessment on the precision of the modelling, as indicated in Ofwat's opinion on Pennon's acquisition of SES Water.[26]

CIs were also assessed by the CMA in its PR19 redeterminations to determine whether the strengthening of the industry cost benchmark from a UQ to the fourth-ranked company for WW base costs, or to the third-ranked company for wholesale wastewater base costs, between the PR19 DD and FD was justified.[27]

A1.5.2 Stochastic frontier analysis methodology

SFA is a well-known benchmarking tool that provides a data-driven assessment of the amount of noise in the models, it does not require the

[25] The operational performance is assessed based on input from YWS and regulatory precedent.
[26] Ofwat (2024), 'Ofwat's Opinion on Pennon's acquisition of SES Water', p. 11.
[27] Competition & Markets Authority (2021), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations. Final report', March, paras 4.442–4.468.

same reliance on previous regulatory decisions as the confidence interval analysis. It has been used by regulators across Europe,[28] and it has also been considered by UK regulators to assess the level of uncertainty in models.[29] For instance, the CMA considered it as part of the PR19 redetermination, although it raised concerns about the focus on pooled SFA rather than on panel SFA. To mitigate the limitations highlighted by the CMA, we have explored both pooled and panel SFA.

A1.6 Post-modelling adjustments

Alongside the base cost modelling, companies also have the opportunity to submit cost adjustment claims (CACs) where the models do not adequately capture all relevant drivers of cost.

At the DD, Ofwat has allowed for sector-wide post-modelling base cost adjustments on: (i) energy price uplifts; (ii) water mains replacement to maintain asset health; (iii) meter replacements; (iv) phosphorus removal; (v) net zero targets. We have assessed these post-modelling adjustments in a separate report, and the efficient cost predictions detailed in this report exclude such adjustments.[30]

As well as sector-wide adjustments, we have also assessed YWS's specific CACs in a separate report.[31] Again, the efficient predictions detailed in this report exclude such adjustments.

[28] See, for example, Bundesnetzagentur (2018), 'Decision BK4-18-056', November; where the German energy regulator employed SFA as the main method used to estimate distribution system operators' efficiency. Moreover, SFA is also used by several other regulators across sectors alongside other methodologies.

[29] See, for example, Office of Rail and Road (2013), 'PR13 Efficiency Benchmarking of Network Rail using LICB', August; Office of Rail and Road (2017), 'Benchmarking regional maintenance costs on England's Strategic Road Network'.

[30] See Oxera (2024), 'Cost adjustment claims', August.

[31] See Oxera (2024), 'Cost adjustment claims', August.

A2 Wholesale water

The following section provides an assessment of YWS's efficient base cost requirements in WW.

A2.1 Model specification

Ofwat has made no changes to its modelling suite at the DD in WW relative to the PR24 base cost modelling consultation. That is, the following cost drivers are accounted for in Ofwat's DD water models.

- **Scale.** Ofwat accounts for scale through connected properties in WRP and WW, and through length of mains in TWD.
- **Treatment complexity**. The proportion of total water treated at complexity bands 3 to 6 and the WAC to account for treatment complexity in WRP and WW.
- **Topography**. In TWD and WW, booster pumping stations per length of main and average pumping head (APH) account for the cost pressures associated with topography within each supply region.
- **Population density.** Ofwat has continued to use the three measures of population density: (i) weighted average density—LAD from MSOA; (ii) weighted average density—MSOA; (iii) properties per length of mains.

While the modelling suite remains unchanged from the PR24 base cost modelling consultation, Ofwat has implemented a 50/50 weighting between models that include weighted average density and properties per length of mains as its density driver in its triangulation approach. For example, in WRP, the four models that incorporate a weighted average density measure have a combined weight of 50% (25% for each weighted average density measure), while the two models that incorporate properties per length of mains also have a combined weight of 50%.

Despite the changes in the weights of the triangulation approach, since the modelling suite remains unchanged, many of the improvements to the modelling that we proposed in our submission alongside YWS's business plan remain relevant for the DD.[32] In light of the additional data and refinements to Ofwat's triangulation approach, we consider that

[32] Oxera (2023), 'An assessment of Yorkshire Water Service's base cost requirements', September.

the following changes to the models would result in a more robust assessment:

- incorporation of properties in treated water distribution (TWD) alongside length of mains as a scale driver;
- WAC to be modelled in levels instead of its natural logarithm;
- combination of booster pumping stations per length of main and APH in the same models.

Under the current cost assessment dataset, with the incorporation of APR data for 2023/24 shared by YWS, these changes result in an improvement to Ofwat's DD models from an operational and/or statistical perspective.[33] We assess each proposal in more detail below.

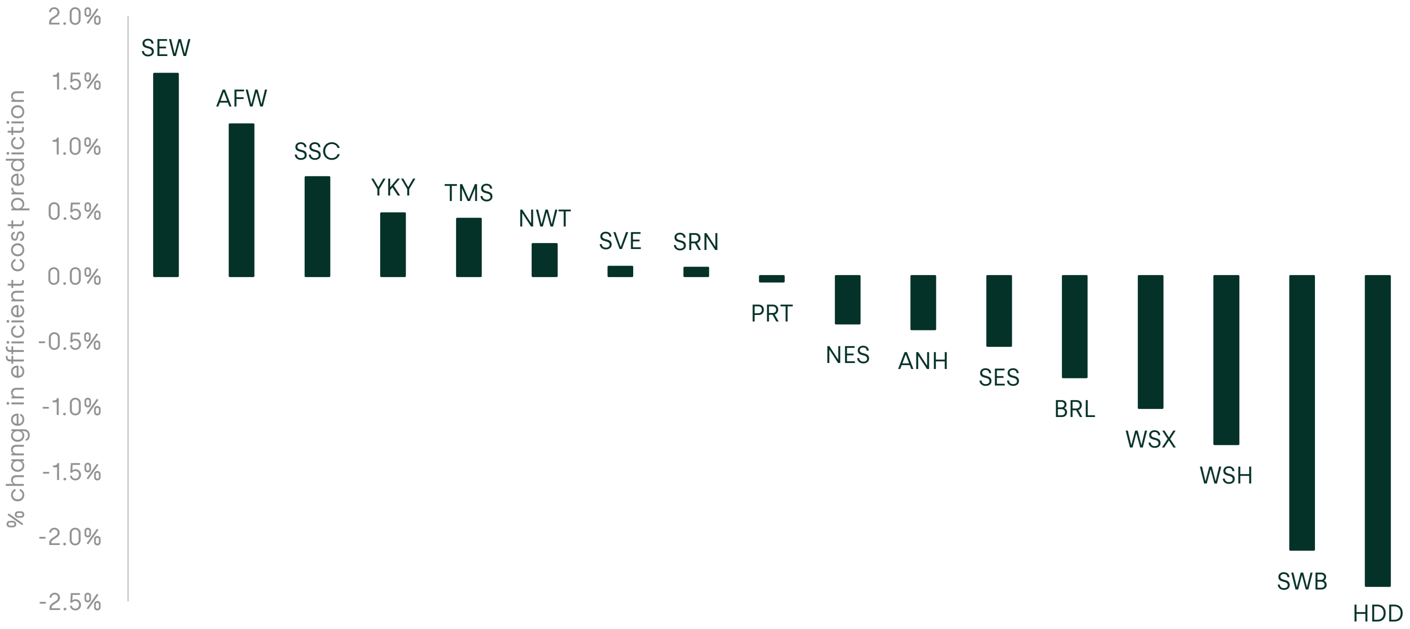
Properties in treated water distribution

Ofwat uses the length of mains as its sole driver of scale in its DD TWD models. We consider that length of mains is a relevant measure of scale in the TWD models from an engineering and statistical perspective. Since costs associated with population growth, such as network reinforcement, are included as part of TWD modelled cost, properties would be a more viable driver to explain such differences in costs between companies. Therefore, we consider length of mains as the scale driver in half of our proposed TWD models, with the remaining half controlling for connected properties.

Ofwat did not include connected properties in its TWD models, arguing that it would have an immaterial impact on companies' outcomes. The following figure summarises the changes to the efficient cost predictions of companies (post frontier shift/RPE) when properties is used as the sole driver of TWD costs under Ofwat's PR24 DD models.

[33] Operational performance is assessed based on views expressed by water companies on Ofwat's proposals, regulatory precedent and input from YWS.

Figure A2.1 Impact of controlling for connected properties in TWD



Note: This analysis is based on a comparison of the efficient cost predictions in Ofwat's TWD models to equivalent models that control for connected properties as the scale variable.
Source: Oxera analysis.

The figure shows that the effect on a company's overall efficient cost prediction when implementing TWD models with properties as the sole driver of scale ranges between c. -2.4% and +1.6%. Therefore, using a similar materiality threshold set by Ofwat for cost adjustment claims, the impact of properties is material for some companies.

In terms of model performance, **connected properties is statistically significant at the 1% level as expected across all models with the expected sign**. The magnitude of the coefficient is also consistent with that of the WRP and WW models. **Controlling for connected properties leads to an improved model fit in two out of six TWD models, with one model's R-squared increasing by one percentage point**. The changes in others model fits are not material.

Given that connected properties works at least as well as length of mains in the models, there is a strong engineering rationale for its inclusion, and its inclusion leads to a material impact for some companies' allowances, we consider that connected properties should be included in half of the TWD models. This would mitigate any bias associated with selecting one driver over another.

Treatment complexity

Ofwat continues to model its WAC measure, which is a weighted proportion variable, in logarithms as opposed to levels. The driver is a weighted sum of proportion variables, making it similar to a typical proportion variable such as the 'water treated at complexity levels 3–6' measure. The primary difference between the two is that the WAC ranges from 1 to 7 while the other complexity driver ranges from 0 to 1. However, this difference is superficial and can be corrected by re-normalising the WAC variable to between 0 and 1 without affecting the estimated relationship (i.e. the coefficient) between complexity and costs, or model performance.

While the CMA argued at the PR19 redetermination that the WAC variable could be modelled in logarithms, it did not interrogate the issue in detail—the CMA's argumentation on this issue is directly counter to a similar decision on how to model proportion variables at PR14, and is counter to operational and economic expectations.

To be clear, when a proportion variable is modelled in levels, the coefficient can be interpreted as the cost impact of increasing the proportion by one percentage point: the cost impact of increasing treatment complexity from 1% to 2% is (approximately) the same as increasing treatment complexity from 50% to 51%. Meanwhile, if the cost driver is modelled in logarithms, the cost impact of increasing treatment complexity by one percentage point varies depending on the current level of treatment complexity. For example, increasing treatment complexity from 1% to 2% would have the same cost impact as increasing treatment complexity from 50% to 100%. **At the PR14 redetermination the CMA considered that such a relationship was operationally and economically unintuitive**.

As noted above, the WAC variable is for all intents and purposes a proportion variable. Therefore, following the operational and economic logic above, the variable should be modelled in levels. Indeed, if this variable is modelled in levels, the coefficient on WAC has a relatively clear interpretation. The change in predicted costs resulting from a shift of 1% of water treated from complexity band 'x' to complexity band 'y' would (approximately) be the estimated coefficient multiplied by (y - x). The magnitude of this coefficient can then be assessed against the expected operational relationship between efficient expenditure and the level of treatment complexity. When the coefficient is modelled in logarithms (as Ofwat currently does), the interpretability of the coefficient is less clear—indeed, neither Ofwat nor the CMA has presented justification for this approach.

The driver in levels also continues to statistically outperform equivalent models that account for the driver in its natural logarithm form. The estimated coefficient on WAC in levels in our proposed WW models are always significant at least to the 5% level, unlike Ofwat's DD models where the driver is insignificant in one WW model, and is of lower statistical significance in general. Furthermore, the proposed WRP models only differ from Ofwat's DD models on the basis of WAC being measured in levels. As a direct comparison, these proposed models have a higher model fit based on R-squared, and the driver is statistically more significant than when it is measured in its natural logarithm form.

We note that Ofwat has not engaged with this issue at the DD. Instead, when discussing companies' critiques of modelling WAC, it focuses on the weights used to construct the measure rather than how the WAC is included in the models. For example, in the PR24 modelling consultation, some companies demonstrated that alternative weights (such as logarithmically increasing weights rather than linearly increasing weights) performed just as well in the models and, as such, Ofwat should undertake a more thorough assessment of whether the weights on each complexity level are aligned with operational expectations. Ofwat has not undertaken this assessment at the DD, which suggests that the weights remain unvalidated.

Topography

Multiple companies proposed including the two drivers of topography (booster pumping stations per length of mains and APH) in the same models at the PR24 modelling consultation and in their business plan submissions. The arguments for including both drivers in the same model included: (i) the drivers being statistically uncorrelated with each other; (ii) arguments from an engineering and operational perspective that they account for different aspects of topography and network complexity; (iii) the drivers perform well together in the same model, can address any omitted variable bias (OVB) and increase model performance.[34]

Ofwat has continued to model the two topography drivers separately at the DD. Ofwat stated three main reasons for continuing to model the two drivers separately, as follows:

[34] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances – Base cost modelling decision appendix', July, p. 21.

1 Booster pumping stations is an imperfect proxy for pumping requirements.
2 Triangulating across different models is simpler and better mitigates omitted variable bias (OVB).
3 Including both drivers in the same model leads to unintuitive outcomes for some companies.

These are discussed in more detail below.

Imperfection of booster pumping stations

Ofwat argues that booster pumping stations is not a measure of network complexity; rather, it is an imperfect measure of companies' pumping requirements. If booster pumping stations was a measure of network complexity, then it would be clearly appropriate to include booster pumping stations and APH in the same model, given that they would be capturing different characteristics (as highlighted by some companies), assuming that the models remained robust.

Ofwat has changed its position on booster pumping stations in the DD relative to previous decisions. For example, Ofwat has referred to booster pumping stations as a 'network complexity' driver in the PR19 econometric modelling consultation, where it was considered to be a proxy for network complexity alongside other network complexity drivers (such as the number of water towers).[35] If booster pumping stations capture some costs associated with network complexity (even if imperfectly),[36] then the two drivers capture different things meaning that both drivers can be included in the same model, even under Ofwat's logic.

We note that booster pumping stations and APH are not strongly correlated with each other (correlation coefficient c. 0.27). This is

[35] Ofwat (2018), 'Cost assessment for PR19: a consultation on econometric cost modelling', March, Table 3.
[36] We note that Ofwat states that booster pumping stations implicitly captures some costs associated with network complexity through its comparison to other network complexity drivers. Ofwat suggests '[n]etwork configuration is complex, and focusing on the number of boosters ignores **other** aspects of network complexity such as service reservoirs, water towers, and the degree of interconnectivity within a network' [emphasis added]. See Ofwat (2024), 'PR24 draft determinations: Expenditure allowances - Base cost modelling decision appendix', July. The observation that booster pumping stations can be grouped into either 'pumping requirements' (in which case it could be interpreted as a 'substitute' for APH) or 'network complexity' (as outlined in this quote) is one of the key concerns with this approach to modelling (where cost drivers are grouped into categories and only one driver from each category is selected), as it can result in the arbitrary omission of relevant drivers.

unexpected if these cost drivers are intended to capture precisely (or even largely) the same operational characteristics.

In its water models, there are two other areas in which Ofwat considers that some cost drivers are 'substitutes' for each other: (i) density (two measures of weighted average density and properties per length of mains); (ii) water treatment complexity (weighted average complexity and the proportion of water treated in complexity bands 3–6). In these cases, the drivers are strongly correlated with each other. For example, the correlation between the two complexity drivers is c. 0.91, while the weakest correlation between the density drivers is c. 0.93. Given that there is a strong conceptual case that these drivers are capturing similar costs, and the drivers are empirically correlated with each other, it may be reasonable to triangulate across models that have different density and complexity drivers (assuming that the models are robust). However, in the case of booster pumping stations and APH, there are mixed views across the industry as to whether APH and booster pumping stations capture different characteristics from an operational perspective, and the two drivers are not empirically correlated with each other, which suggests that the two drivers may capture different cost pressures.

Moreover, the two drivers can be included in the same model, even if they are assumed to capture similar characteristic. Assuming that the two drivers are not perfectly collinear and that each cost driver captures the relevant characteristic imperfectly, controlling for any one driver over another will result in a bias (issues pertaining to the bias are discussed in more detail below). We note that Ofwat has used models that control for several related cost drivers to set allowances at previous determinations,[37] including at PR19 when Ofwat controlled for population density[38] and STW-size in its BR models.[39]

The CMA explored this issue in the PR19 redetermination. It stated:[40]

We recognised that excluding APH risked causing omitted variable bias. However, **excluding booster pumping stations and including APH would**

[37] Indeed, controlling for population density and its squared term could be seen as an example of using two cost drivers.
[38] Ofwat argues that, in the BR models, higher density 'may allow for the use of larger, more efficient, treatment works'. See Ofwat (2019), 'Supplementary technical appendix: Econometric approach', January, p. 22.
[39] See Ofwat (2019), 'PR19 slow track draft determinations: Securing cost efficiency technical appendix', December, Table A2.2.
[40] CMA (2021), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations: Final report', March, para. 4.81.

also create omitted variable bias. Including both would not be appropriate as APH would not be statistically significant in the WW1 and WW2 models. Therefore, in this case, we did not consider omitted variable bias was a substantial enough reason to use APH in the base cost models. [emphasis added]

There are two relevant insights from this decision. First, **the CMA acknowledged explicitly that a model that controls for APH and not booster pumping stations and vice versa (i.e. Ofwat's approach at the DD) risked generating OVB**. Second, **the CMA explored controlling for both drivers in the same model (as we present in this report) but found that APH was statistically insignificant**. That is, the sole justification provided for not including both drivers in the same model was that APH was statistically insignificant.[41] This is no longer the case in the majority of the models that we have explored.

Triangulation avoids OVB

Ofwat argues that triangulating across models mitigates OVB. While it is the case that triangulating across models could *reduce* the bias associated with omitted drivers in any one model, it does not *eliminate* it entirely. Ofwat assumes that the bias in the models that control for booster pumping stations is perfectly offset by the bias in the models that control for APH. This assumption is unsupported by the evidence, given that companies' performance in models that control for both drivers (which have no OVB as a result of this issue) is often materially different to their performance in Ofwat's DD models.

That is, Ofwat's argument that APH and booster pumping stations cannot be in the same model because some companies' performance changes materially assumes that the triangulation approach is the unbiased estimate. However, **as noted by the CMA (see above), models that omit relevant drivers of expenditure will suffer from OVB**. A direct test of OVB is to assess the statistical significance and interpretability of the relevant measure directly in the model. This direct test supports the inclusion of both measures in the same model. Ofwat has not provided evidence that triangulating the results from different models—

[41] Note that there were other issues raised regarding the use of APH at the PR19 redetermination, including data quality and its statistical performance in the models. However, these concerns related to the use of APH in general, and not the use of two drivers to capture the same costs. We understand that some of these concerns with APH are less relevant at PR24 (for example, due to improvements in the data quality for APH).

all of which are potentially biased as a result of this issue—perfectly offsets the individual biases in each model.

Ofwat also argues that triangulating across models is more consistent with its approach to developing 'simple' models. This is a value judgement that is based on one interpretation of the word 'simple'. It is unclear how models that control for both drivers are materially more complex than models that control for only one of the drivers. Moreover, by not including both drivers in the same model, Ofwat has essentially doubled the size of its modelling suite, which could be considered more complex than including both drivers in the same model (in other price controls, Ofwat has reduced the number of models).

Following the APR update, APH reduces in statistical significance across the six WW models, and becomes insignificant (albeit still close to the 10% level) in two of Ofwat's WW DD models, while boosters per length remains statistically significant across all models. In our models that control for both drivers, both APH and boosters per length are statistically significant across all models with the exception of three WW models where APH is close to significance at the 10% level. Furthermore, the model fits based on R-squared improve when APH and boosters per length are included in the same model,[42] relative to their equivalent models in Ofwat's DD.[43]

We note that, if there are concerns that either booster pumping stations or APH fail to capture company-specific cost pressures, these could be addressed through the CAC process for specific companies.

A2.2 Benchmark analysis

In this section, we examine the precision of Ofwat's econometric modelling in WW—in particular, we examine: confidence intervals around companies' cost predictions (section A2.2.1), and two SFA approaches (pooled and panel) as alternative benchmarking modelling (section A2.2.2).

A2.2.1 Confidence intervals

At PR19, the CMA did not consider that an industry cost benchmark more stretching than the UQ for WW base costs was justified, and we find

[42] We note that the model fit increases by construction when more cost drivers are added to the model, and could be the result of 'overfitting'. If we examine other measures of model fit that penalise overfitting, such as the Akaike information criterion (AIC), our models continue to outperform the DD models. This suggests that the improved model fit is not a result of overfitting.
[43] We note that the statistical performance of these models is somewhat sensitive to the treatment of PRT's data (see section A1). The data corrections that we adopt in this report will need to be confirmed at the FD stage.

that the width of CIs (i.e. the level of uncertainty) at PR24 has increased since. This indicates that less stretching efficiency targets could be justified at PR24, or at the very least, that there is no evidence supporting a more stringent efficiency target than the UQ.

This is especially true for WRP models where the width of the 95% CIs has significantly increased, from 16% in the CMA's PR19 redeterminations to 21% in Ofwat's PR24 DD. This also holds for WW models with an increase from 10% to 13%. However, the degree of certainty of cost predictions arising from treated water distribution (TWD) models improved slightly, with a width of 13%. This can be observed in Table A2.1 below.

Table A2.1 Estimated CIs in Ofwat's PR24 DD models for WW base costs including the APR24 update

	WRP	TWD	WW
Model 1	19%	14%	14%
Model 2	22%	14%	14%
Model 3	21%	17%	14%
Model 4	23%	12%	14%
Model 5	20%	10%	12%
Model 6	20%	12%	12%
Model 7	-	-	14%
Model 8	-	-	14%
Model 9	-	-	15%
Model 10	-	-	15%
Model 11	-	-	12%
Model 12	-	-	12%
Average (PR24)	**21%**	**13%**	**13%**
Average (PR19)	**16%**	**13%**	**10%**

Note: The figures presented in the table represent the width of the 95% CI around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% CI ranges from -X% of the predicted costs to + X% of the predicted costs. We acknowledge that aggregating and averaging CIs is theoretically problematic, but we provide this figure to improve the readability of the table.
Source: Oxera analysis from Ofwat (2024), 'PR24-DD-Base-costs-water-model-1', June; Ofwat (2024), 'PR24-DD-Base-costs-water-model-2', June; Ofwat (2024), 'Water base

cost (Stata do file)', July; 'APR Industry Datashare 2024 Publish V1', provided by YWS, July.

There is strong evidence that the level of uncertainty in Ofwat's models has materially increased since the PR19 redeterminations, where the CMA applied a UQ benchmark. Given this increase in uncertainty, a less stringent benchmark may be more appropriate (e.g. an upper-tercile benchmark). That is, far from being '[in]sufficiently stretching for some companies',[44] the evidence suggests that the UQ benchmark is more stringent than the models would support.

A2.2.2 Stochastic frontier analysis

Table A2.2 below shows how the average efficiency gap in the pooled SFA models compares to the average efficiency gap at a UQ benchmark, as well as whether there is any statistically significant inefficiency in the sample. These results include the new outturn data from the annual performance report for 2023/24.

[44] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances – Base cost modelling decision appendix', July, p. 23.

Table A2.2 Average estimated efficiency gaps—Ofwat's PR24 DD models including the APR24 update (2019/20–2023/24)

	Average gap to UQ			Average inefficiency SFA (pooled)		
	WRP	**TWD**	**WW**	**WRP**	**TWD**	**WW**
Model 1	17%	11%	11%	18%*	1%	8%
Model 2	19%	8%	11%	9%	0%	8%
Model 3	18%	8%	11%	21%**	4%	9%
Model 4	20%	12%	10%	9%	7%	10%
Model 5	19%	10%	9%	21%***	1%	12%***
Model 6	20%	13%	9%	15%	8%	13%**
Model 7	-	-	14%	-	-	14%**
Model 8	-	-	14%	-	-	11%
Model 9	-	-	15%	-	-	17%***
Model 10	-	-	15%	-	-	15%*
Model 11	-	-	10%	-	-	15%***
Model 12	-	-	12%	-	-	13%**
Average	**19%**	**10%**	**11%**	**15%**	**3%**	**12%**

Note: The last three columns include the likelihood ratio test for the presence of inefficiency in the sample. *, **, and *** show statistical significance at the 10%, 5% and 1% levels, respectively. No asterisk indicates that there is no statistically significant inefficiency in the sample.
Source: Oxera analysis from Ofwat (2024), 'PR24-DD-Base-costs-water-model-1', June; Ofwat (2024), 'PR24-DD-Base-costs-water-model-2', June; Ofwat (2024), 'Water base cost (Stata do file)', July; 'APR Industry Datashare 2024 Publish V1', provided by YWS, July.

The analysis indicates that the average gap to the UQ exceeds the average efficiency estimated in the SFA models for both WRP and TWD. Notably, half of the WRP models and all TWD models fail to detect any statistically significant inefficiency within the sample, suggesting that much or all of the estimated efficiency gap in Ofwat's models is attributable to statistical noise. In the WW models, the average efficiency gap in SFA is roughly equivalent to what a UQ benchmark would indicate. However, about half of the WW models do not identify any statistically significant inefficiency in the sample.

While the pooled SFA models suggest that a UQ may be broadly appropriate in the WW models, and that a less stringent benchmark should be applied in the WRP and TWD models, the panel SFA models do not detect any statistically significant inefficiency in the sample across any model specification. This implies that, after accounting for unobserved company heterogeneity, most (or all) of the remaining estimated efficiency gap is driven by statistical noise. In turn, this would suggest that the UQ benchmark is unsupported by the evidence.

A2.3 YWS's estimated allowance

Table A2.3 below shows our assessment of YWS's efficient base expenditure for AMP8 in WW, as well as a comparison to Ofwat's DD models.

Table A2.3 YWS's efficient cost predictions in WW

	Submitted costs	Efficient modelled cost
PR24 DD models with APR update	2334	1657
Our proposed models with APR update	2334	1691

Note: The estimated values above include a UQ benchmark, FS/RPE assumptions as per Ofwat's DD decision, and are in £m 2022/23 prices. The values prior to the APR update also do not incorporate any changes made to the cost assessment dataset, or in the case of the proposed models, reflect the final proposals.
Source: Oxera analysis.

YWS's efficient cost prediction under the proposed models, which offer an improvement to Ofwat's DD models due to the aforementioned reasons, is c. £1,691m. This is c. £34m higher than that predicted by Ofwat's DD models. That is, using better-specified models that outperform Ofwat's modelling from an economic and statistical perspective increases YWS's estimated efficient cost prediction. However, YWS still has a cost gap of c. £643m, although part of this is accounted for by the base cost modelling adjustments.[45]

[45] For our assessment of YWS's base cost adjustments, please refer to the submission document Oxera (2024), 'Cost adjustment claims', August.

A3 Wholesale wastewater network plus

The following section provides an assessment of YWS's efficient base cost requirements for WWNP.

A3.1 Model specification

Ofwat has removed ten of its 17 PR24 consultation models and included three models for SWC, two models for SWT and two models for WWNP. The following cost drivers are accounted for in Ofwat's DD wastewater models.[46]

- **Scale.** Ofwat accounts for scale through sewer length in SWC and load in SWT and WWNP as they capture the scale of operations for companies.
- **Economies of scale at STWs.** Ofwat adopts two economies of scale drivers: the percentage of load treated in complexity bands 1 to 3, and the WATS in SWT and WWNP. Relative to the PR24 modelling consultation, Ofwat has dropped models that include the percentage of load treated in STWs ≥ 100,000 people.
- **Treatment complexity**. Ofwat continues to use the percentage of load with ammonia permit <= 3mg/l as the treatment complexity driver in its SWT and WWNP models.
- **Network topography.** Ofwat has continued to use the pumping capacity per sewer length as the network topography driver in its SWC and WWNP models.
- **Population density.** Ofwat has continued to use the three measures of population density: (i) weighted average density—LAD from MSOA; (ii) weighted average density—MSOA; (iii) properties per length of mains.
- **Urban rainfall.** Ofwat has only included models that include urban rainfall as a cost driver. Ofwat uses the Urban MSOA rainfall measure, that is, the average rainfall in a company area multiplied by urban company area using rainfall data for middle super output areas (MSOAs), as a driver of costs associated with urban rainfall across its SWC and WWNP models.

As well as the changes to the modelling suite, Ofwat has implemented a 50/50 weighting across models that include properties per sewer length

[46] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances – Base cost modelling decision appendix', July, p. 30.

as its density driver, and the weighted average density measures. For example, across SWC, it has given a 50% weight to the model with properties per sewer length, and 25% each for the two models that account for weighted average density. Furthermore, it has also adopted a 50/50 triangulation over the bottom-up (sum of SWC and SWT) and the top-down (WWNP) approaches.

Despite the positive changes made by Ofwat to its models, many of the modelling changes proposed for WWNP remain unaccounted for. We propose the following change to Ofwat's WWNP modelling suite: dropping the percentage of load treated in bands 1–3 as the economies of scale driver.

Under the current cost assessment dataset, with the APR data incorporated, this modification results in an improvement over Ofwat's SWT and WWNP models.

The WATS variable is operationally superior to the percentage of load treated in size bands 1–3. The DD models include the percentage of load treated in size bands 1–3 as one of its economies of scale drivers in the SWT and WWNP models. However, unlike the WATS driver, this relies on the arbitrary threshold of bands 1–3 which assumes that there is a step change in efficient costs at the STW level. This implies that all STWs in this size band have the same level of efficient unit costs (e.g. STWs in size band 3 cannot benefit from additional economies of scale compared with those in size band 1), and that all STWs in size bands 4 and above have the same level of efficient unit costs (e.g. STWs above size band 5 cannot benefit from additional economies of scale compared with those in size band 4). Conversely, WATS does not rely on any arbitrary thresholds, and allows for a smoother relationship between STW size and efficient costs.

Furthermore, the WATS variable is statistically superior to the percentage of load treated in bands 1–3 variable as the latter is statistically insignificant in the SWT model used by Ofwat. The models with this variable have a lower model fit compared with the model with WATS, thus reaffirming its exclusion from our proposed models.

A3.2 Benchmark analysis

In this section, we examine the precision of Ofwat's econometric modelling in wholesale wastewater. We examine: confidence intervals around companies' cost predictions (section A3.1.1), and two SFA approaches (pooled and panel) as alternative benchmarking modelling (section A3.1.2).

A3.1.1 Confidence intervals

Similar to WW, our assessment is that Ofwat's choice of setting a benchmark to the UQ level of efficiency seems to correspond to the most challenging benchmark choice that Ofwat could have reasonably assumed at the DD stage.

Table A3.1 Estimated CIs in Ofwat's PR24 DD models for wholesale wastewater base costs including the APR24 update

	SWC	SWT	WWNP
Model 1	12%	16%	8%
Model 2	12%	10%	7%
Model 3	13%	-	-
Average (PR24)	**12%**	**13%**	**8%**
Average (PR19)	**14%**	**14.5%**	**NA**

Note: The figures presented in the table represent the width of the 95% CI around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% CI ranges from -X% of the predicted costs to +X% of the predicted costs. We acknowledge that aggregating and averaging CIs is theoretically problematic, but we provide this figure to improve the readability of the table.
Source: Oxera analysis from Ofwat (2024), 'PR24-DD-Base-costs-wastewater-model-1', June; Ofwat (2024), 'PR24-DD-Base-costs-wastewater-model-2', June; Ofwat (2024), 'Wastewater network plus base cost (Stata do file)', July; 'APR Industry Datashare 2024 Publish V1', provided by YWS, July.

The table shows that the confidence intervals in the PR24 models for SWC and SWT are slightly narrower than the confidence intervals at the PR19 redetermination. Nevertheless, these continue to be wider than for WW at the PR19 redetermination, where the CMA applied a UQ as benchmark.

The main exception to this is in the WWNP models, where the confidence intervals are narrower on average than in the PR19 models, indicating that the WWNP models estimate companies' costs with a lower degree of uncertainty than the bottom-up models.

A3.1.2 Stochastic frontier analysis

Table A3.2 below shows how the average efficiency gap in the pooled SFA models compares to the average efficiency gap at a UQ benchmark, as well as whether there is any statistically significant inefficiency in the sample. These results include the new outturn data from the annual performance report for 2023/24.

Table A3.2 Average estimated efficiency gaps—Ofwat's PR24 DD models including the APR24 update (2019/20–2023/24)

	Average gap to UQ			Average inefficiency SFA (pooled)		
	SWC	SWT	WWNP	SWC	SWT	WWNP
Model 1	4%	11%	4%	6%	10%***	8%***
Model 2	5%	9%	4%	7%	7%***	6%***
Model 3	7%	-	-	5%	-	-
Average	**5%**	**10%**	**4%**	**6%**	**10%**	**7%**

Note: The last three columns include the likelihood ratio test for the presence of inefficiency in the sample. *, **, and *** show statistical significance at the 10%, 5% and 1% levels, respectively. No asterisk indicates that there is no statistically significant inefficiency in the sample.
Source: Oxera analysis from Ofwat (2024), 'PR24-DD-Base-costs-wastewater-model-1', June; Ofwat (2024), 'PR24-DD-Base-costs-watsewater-model-2', June; Ofwat (2024), 'Wastewater network plus base cost (Stata do file)', July; 'APR Industry Datashare 2024 Publish V1', provided by YWS, July.

The analysis indicates that the average gap to the UQ is aligned with the average efficiency estimated in the SFA models for all three aggregations. Notably, for the models in SWT and WWNP, the pooled SFA is able to find statistically significant inefficiencies across companies, but these are not significant for any of the three SWC models. Moreover, e note that while the pooled SFA models suggest that a UQ may be broadly appropriate in the wholesale wastewater, the results from the panel SFA models do not detect any statistically significant inefficiency in the sample across any model specification. This implies that, after accounting for unobserved company heterogeneity, most (or all) of the remaining estimated efficiency gap is driven by statistical noise.

A3.3 Yorkshire Water's estimated allowance

Table A3.3 below shows our assessment of YWS's efficient base expenditure for AMP8 in WWNP, as well as a comparison to Ofwat's DD models.

Table A3.3 YWS's efficient cost predictions in WWNP

	Submitted costs	Efficient modelled cost
PR24 DD models with APR update	1,723	1,752
Our proposed models with APR update	1,723	1,752

Note: The estimated values above include a UQ benchmark, FS/RPE assumptions as per Ofwat's DD decision, and are in £m 2022/23 prices. The values prior to the APR update

also do not incorporate any changes made to the cost assessment dataset, or in the case of the proposed models, reflect the final proposals.
Source: Oxera analysis.

YWS's efficient cost prediction under the proposed models is c. £1,752.1m, which is marginally higher than that predicted by Ofwat's DD models. By including only the better performing and operationally superior cost drivers in the proposed models, YWS's allowance increases post the APR update, which is material after accounting for the base cost adjustments.

A4 Bioresources

The following section presents an assessment of YWS's bioresources base cost expenditure requirements.

A4.1 Model specification

Ofwat has included four unit cost models, dropping the six total cost models presented in its PR24 consultation. Thus, its PR24 modelling approach is slightly different from that of PR19 for bioresources.

It states the following reasons for solely using the unit cost models: (i) unit cost bioresources models omit the sludge-produced scale variable, thus imposing a constant returns to scale assumption; (ii) unit cost models are consistent with bioresources average revenue control and will support the bioresources market; (iii) the population density variables and total volume of sludge produced are statistically insignificant in the total cost models.

The cost drivers of these unit cost models are economies of scale in sludge treatment, and location of STWs relative to sludge treatment centres. Ofwat suggests that both of these variables are 'somewhat under company control'[47], thereby using population density and the size of STWs as proxies to capture similar effects. Thus, the following cost drivers are included in the unit cost models for analysis:

- percentage of load treated in bands 1–3;
- weighted average density—LAD from MSOA (log);
- weighted average density—MSOA (log);
- number of STWs per property (log).

Since these variables are highly correlated, they capture similar information. Ofwat triangulates over the four models by giving equal weight to all to produce the efficient cost predictions for the companies.

A4.2 Benchmark analysis

In this section, we examine the precision of Ofwat's econometric modelling in wholesale wastewater. We examine confidence intervals around companies' cost predictions (section A4.1.1), and two SFA

[47] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances – Base cost modelling decision appendix', July, p. 55.

approaches (pooled and panel) as alternative benchmarking modelling (section A4.1.2).

A4.1.1 Confidence intervals

Table A4.1 below shows the average confidence intervals around companies' cost predictions in the bioresources base models.

Table A4.1 Estimated CIs in Ofwat's DD models for BR including the APR24 update

	Bioresources unit cost models
Model 1	14%
Model 2	19%
Model 3	20%
Model 4	20%
Average (PR24)	**18%**
Average (PR19)	**21%**

Note: The figures presented in the table represent the width of the 95% CI around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% CI ranges from -X% of the predicted costs to +X% of the predicted costs. We acknowledge that aggregating and averaging CIs is theoretically problematic, but we provide this figure to improve the readability of the table.
Source: Oxera analysis from Ofwat (2024), 'PR24-DD-Base-costs-wastewater-model-1', June; Ofwat (2024), 'PR24-DD-Base-costs-wastewater-model-2', June; Ofwat (2024), 'Bioresources base cost (Stata do file)', July; 'APR Industry Datashare 2024 Publish V1', provided by YWS, July.

The confidence intervals from the bioresources models are slightly narrower than at the PR19 redeterminations. Nevertheless, these are still typically wider than in other services where the CMA applied a UQ benchmark. These relatively wide confidence intervals suggest that Ofwat should decrease the stringency applied in this service, as there is high uncertainty around the companies cost estimations.

A4.1.2 Stochastic frontier analysis

Table A4.2 below shows how the average efficiency gap in the pooled SFA models compares to the average efficiency gap at a UQ benchmark, as well as whether there is any statistically significant inefficiency in the sample. These results include the new outturn data from the annual performance report for 2023/24.

Table A4.2 Average estimated efficiency gaps—Ofwat's PR24 DD models including the APR24 update FD models (2019/20–2023/24)

	Average gap to UQ	Average inefficiency SFA (pooled)
Model 1	18%	0%
Model 2	24%	0%
Model 3	24%	0%
Model 4	26%	0%
Average	**23%**	**0%**

Note: The last three columns include the likelihood ratio test for the presence of inefficiency in the sample. *, **, and *** show statistical significance at the 10%, 5% and 1% levels, respectively. No asterisk indicates that there is no statistically significant inefficiency in the sample.
Source: Oxera analysis from Ofwat (2024), 'PR24-DD-Base-costs-water-model-1', June; Ofwat (2024), 'PR24-DD-Base-costs-water-model-2', June; Ofwat (2024), 'Water base cost (Stata do file)', July; 'APR Industry Datashare 2024 Publish V1', provided by YWS, July.

The table shows that the average efficiency gap to the UQ is 23%, which is materially higher than in other services. However, the SFA models suggest that all of this estimated efficiency gap is due to statistical noise. This further supports the position that a UQ benchmark is not appropriate in BR, and other benchmarks (such as median) should be selected.

We note that the results from the panel SFA models do not detect any statistically significant inefficiency in the sample across any model specification, akin to the pooled SFA models.

A4.3 YWS's estimated allowance

Table A4.3 shows an assessment of YWS's efficient base cost expenditure for AMP8 in bioresources.

Table A4.3 YWS's efficient cost predictions in bioresources

	Submitted costs	Efficient modelled cost
PR24 DD models with APR update	338	359

Note: The estimated values above include a UQ benchmark, FS/RPE assumptions as per Ofwat's DD decision, and are in £m 2022/23 prices. The values prior to the APR update also do not incorporate any changes made to the cost assessment dataset or, in the case of the proposed models, reflect the final proposals.
Source: Oxera analysis.

YWS's efficient cost allowance under Ofwat's models is c. £359m, which is higher than its submitted costs by c. £21m. However, the allowances have fallen post the APR update by c. £10m, owing to an increase in benchmark stringency from 0.91 to 0.88. The fall in cost allowances post the APR update also appear to be driven by the outturn period (2023/24) having a unit cost significantly lower than the mean unit costs for the modelled period before the update (2019–2023) for YWS. Notably, the average unit costs for the modelled period fall from £0.62/tds for 2019–2023 to £0.52/tds for 2020–2024.

The following are plausible explanations for a lower cost allowance after the update: (i) an increase in the weighted average density—LAD from MSOA measure in the outturn period 2023/24; (ii) an increase in the weighted average density—MSOA measure in the outturn period 2023/24; (iii) a decrease in the number of STWs per property.

Nonetheless, YWS's efficient cost predictions for bioresources are materially greater than its submitted costs, suggesting that YWS performs well under Ofwat's bioresources models.

A5 Residential retail

The following section provides an assessment of YWS's efficient base expenditure requirements for residential retail.

A5.1 Model specification

Ofwat has made minor changes to its modelling suite at the DD relative to the PR24 base cost modelling consultation. The final suite of Ofwat models is as follows: (i) two bad debt cost (RDC) models; (ii) two other cost (ROC) models; (iii) four total retail cost (RTC) models. In all the models, the dependent variable is specified on a cost-per-household basis.

The results of these models are then combined to compute the final retail base cost allowance. Ofwat differentiates between disaggregated/bottom-up (ROC + RDC) and aggregated/top-down (RTC) retail cost models. Ofwat applies a higher weight of 75% to the top-down models and 25% to the bottom-up models.

Ofwat focused on the same key cost drivers as in the base cost modelling consultation, which are as follows.

- Economies of scale, through the inclusion of connected properties as a scale variable in one of its other cost (ROC) and three of its total cost (RTC) models.
- Revenue at risk, through average bill size in all of its bad debt (RDC) and RTC models.
- Propensity to default, through measures of deprivation (percentage of households with payment default, income deprivation score) in its RDC and RTC models.
- Type of customer, through the proportion of dual-service households in one of its ROC models.
- The effect of the COVID-19 pandemic, through the time dummies for the year 2020 and 2021 in all of its RDC and RTC models.

The primary difference between the DD models and the models that Ofwat presented in the PR24 cost modelling consultation is that Ofwat now excludes models that accounted for 'average number of county court judgements' as a measure of deprivation. Ofwat found that the cost driver was not statistically significant and performed worse in the models than other deprivation measures, an argument that was supported by companies in their responses to the modelling consultation. Consequently, models including this variable were also removed, reducing the total number of models from eleven (at the modelling consultation) to eight (at the DD).

Given that the models are largely unchanged, the improvements to the model specification that we outlined in our submission alongside YWS's business plan remain relevant. Specifically, Ofwat's treatment of the COVID-19 years (i.e. including time dummies) is a crude tool for addressing the hypothesised issues with modelling in this period, and alternative methods are more robust. This is discussed in more detail below.

Does COVID-19 require special treatment?

At the PR24 modelling consultation, Ofwat found that the performance of the PR19 models deteriorated materially when they were updated to include data for the first years of AMP7. Specifically, the model fit worsened and the coefficients on relevant cost drivers (such as deprivation) became statistically insignificant and sometimes counterintuitive. Ofwat argued that this was due to an increase in bad-debt provisions during these years that were not explained by the cost drivers and, therefore, include COVID-19 dummies in the RDC and RTC models to capture this 'unexplained' increase in bad debt provisions.

The issues with the use of time dummies to capture the impact of COVID-19 is explained in more detail below. However, we note that the COVID-19 dummies are no longer required to estimate models that have directionally intuitive, statistically significant (or close to) coefficients and reasonably high model fit. Therefore, we consider that the case for including COVID-19 dummies is significantly weakened relative to the PR24 modelling consultation.

The only area in which removing the COVID-19 dummies leads to a perceived deterioration in model quality is with respect to the RESET test (Ofwat's test for model specification). Ofwat's models (that include the COVID-19 dummies) pass the RESET test, while the models fail the RESET test when the dummies are excluded from the specification. A simplistic reading of this result could suggest that the models are misspecified when the COVID-19 dummies are omitted from the models. However, this interpretation of the tests is incorrect when the issue is investigated in more detail.

Table A5.1 shows the p-value of the RESET test in alternative retail models, estimated over different time periods and samples of companies.

Table A5.1 P-values of RESET tests across sensitivities

Sample companies	Sample years	COVID-19 dummies	RDC1	RDC2	RTC1	RTC2	RTC3	RTC4
All companies	2014–24	Yes	0.62	0.25	0.15	0.57	0.32	0.54
All companies	2014–24	No	0.07	0.03	0.03	0.02	0.00	0.00
Excluding NWT and WSH	2014–24	No	0.34	0.36	0.85	0.65	0.15	0.25
All companies	2014–19	No	0.36	0.01	0.01	0.02	0.00	0.00
Excluding NWT and WSH	2014–19	No	0.62	0.25	0.15	0.57	0.32	0.54

Source: Oxera analysis.

The table shows that Ofwat's models generally fail the RESET test when the models are estimated over 2014–19 (the time period used at PR19). That is, the models failing the RESET test is not an issue relating to the inclusion or exclusion of COVID-19 dummies; rather, it appears to be a general issue with Ofwat's model specification. Note that when two companies (NWT and WSH) are excluded from the models, the models pass the RESET test without COVID-19 dummies in both of the time periods that we explored.

A reasonable interpretation of this finding is that the inclusion of COVID-19 dummies, rather than effectively controlling for the spike in doubtful debts during the COVID-19 years, is instead capturing outliers that the RESET test may mistake for non-linearities. Therefore, we do not consider that the fact that the retail models fail the RESET test when COVID-19 dummies are omitted should be determining factor.

Are time dummies the best method of accounting for COVID-19?

Assuming that COVID-19 remains a modelling issue, the use of time dummies is not an appropriate means of accounting for this. Time dummies are blunt instruments that capture all of the cost pressures in a given year, do not isolate the impact of COVID-19 on costs, and artificially depress forward-looking allowances.

Ofwat argues that COVID-19 resulted in an increase in doubtful debt provisions, which would lead to an increase in bad debt costs and (by extension) total costs. If the time dummies purely captured this effect, we would expect the time dummies to be statistically insignificant in the ROC models, given that 'other costs' does not include doubtful debt provisions. However, the time dummy for 2020 is positive and statistically significant, indicating that there are other cost pressures in this year that are unrelated to COVID-19 and doubtful debt. Removing these cost pressures through time dummies is inappropriate in this context and could result in underfunding companies in AMP8. Doing so would assume that there will be no forward-looking cost pressures in AMP8 that are similar to the effect of COVID-19 (e.g. other macroeconomic shocks), which is (at best) an untested assumption by Ofwat.

There are several more targeted approaches that could be adopted to address the impact of COVID-19 on doubtful debt provisions. We have explored the following:

- smoothing doubtful debt costs over the modelling period, in line with Ofwat's treatment of depreciation;
- using linear interpolation to impute the doubtful debt costs in 2020 and 2021;
- imputing the doubtful debt costs in 2020 and 2021 using the average doubtful debt costs in 2017–19.

In addition to avoiding strong and untested forward-looking assumptions, these models perform well against Ofwat's modelling criteria. All of the estimated coefficients in these models are statistically significant, and they retain the expected sign. In terms of model fit, the adjusted R-square is higher or comparable to Ofwat's. That is, these models represent a clear improvement over Ofwat's DD models. A more detailed summary of these models can be found in section A6.4.

In line with our previous proposals, we consider that smoothing doubtful debt costs over the modelling period is the preferred approach based on our comparison of model performance. However, our alternative approaches are also robust and can be considered viable options depending on specific circumstances or preferences.

A5.2 Benchmark analysis

In this section, we examine the precision of Ofwat's econometric modelling—in particular, we examine: confidence intervals around companies' cost predictions, and two SFA approaches (pooled and panel) as alternative benchmarking modelling.

A5.2.1 Confidence intervals

In contrast with WW, the CMA did not make a decision on residential retail as it was not part of its PR19 redeterminations. Instead, we compare the width of the 95% CI around companies' cost predictions in Ofwat's PR19 FD and PR24 models. Table A5.2 includes the average uncertainty associated with the companies' cost predictions for each of the PR24 models. The last two rows refer to the average uncertainty across the suite of models for each of the approaches for both PR24 and PR19.[48]

Table A5.2 Estimated CIs in Ofwat's PR24 DD models for residential retail including the APR24 update

	Bad debt costs	Other costs	Total costs
Model 1	18%	10%	9%
Model 2	19%	10%	12%
Model 3	-	-	9%
Model 4	-	-	12%
Average—FD24	**19%**	**10%**	**11%**
Average—FD19	19%	12%	13%

Note: The figures presented in the table represent the width of the 95% CI around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% CI ranges from -X% of the predicted costs to +X% of the predicted costs. We acknowledge that aggregating and averaging CIs is theoretically problematic, but we provide this figure to improve the readability of the table.
Source: Oxera analysis from Ofwat (2024), 'PR24-DD-Base-costs-residential-retail-1', June; Ofwat (2024), 'PR24-DD-Base-costs-residential-retail-2', June; Ofwat (2024), 'Residential retail (Stata do file)', June; 'APR Industry Datashare 2024 Publish V1', provided by YWS, July.

The results show that bad debt costs models for PR19 produce CIs with a width of 19%, which align with the resulting average widths for PR24. This indicates that the degree of certainty in the predictions is roughly

[48] The CI associated with the PR19 models include only the data that was available at the time of the FD—i.e. up and including financial year 2018/19.

the same despite the increased number of observations. However, the certainty of the predictions has increased marginally for other costs and total costs models from PR19 to PR24, with the width of the CIs decreasing from 12% to 10% and from 13% to 11%, respectively.

A5.2.2 Stochastic frontier analysis

Table A5.3 below shows how the average efficiency gap in the pooled SFA models compares to the average efficiency gap at a UQ benchmark in the DD models for residential retail. These results include the new outturn data from the annual performance report for 2023/24.

Table A5.3 Average estimated efficiency gaps—Ofwat's PR24 DD models including the APR24 update (2019/20–2023/24)

	Average gap to UQ			Average inefficiency SFA (pooled)		
	Bad debt costs	**Other costs**	**Total costs**	**Bad debt costs**	**Other costs**	**Total costs**
Model 1	22%	11%	9%	0%	7%	10%***
Model 2	21%	10%	7%	0%	6%	10%***
Model 3	-	-	8%	-	-	11%***
Model 4	-	-	7%	-	-	10%***
Average	**22%**	**10%**	**8%**	**0%**	**7%**	**10%**

Note: The last three columns include the likelihood ratio test for the presence of inefficiency in the sample. *, **, and *** show statistical significance at the 10%, 5% and 1% levels, respectively. No asterisk indicates that there is no statistically significant inefficiency in the sample.
Source: Oxera analysis from Oxera analysis from Ofwat (2024), 'PR24-DD-Base-costs-residential-retail-1', June; Ofwat (2024), 'PR24-DD-Base-costs-residential-retail-2', June; Ofwat (2024), 'Residential retail (Stata do file)', June; 'APR Industry Datashare 2024 Publish V1', provided by YWS, July.

The table shows that no bad debt cost model detects any statistically significant inefficiency, further supporting the observation that these models estimate costs with a high degree of uncertainty. Similarly, although the other cost models are able to capture some inefficiency, this is not statistically significant. Meanwhile, total cost models detect statistically significant inefficiency, and the estimated efficiency gap in these SFA models is comparable to (or greater than) that implied by the UQ benchmark. However, the panel SFA models do not detect any statistically significant inefficiency, suggesting that much of the estimated efficiency gap may be driven by statistical noise.

The evidence regarding the most appropriate benchmark in the retail models is mixed: while the CI analysis and the pooled SFA models may support a UQ benchmark in the other and total cost models, the bad debt cost models are estimated with materially higher uncertainty, and the panel SFA models suggest that the estimated efficiency gap in all models is driven by statistical noise rather than inefficiency.

A5.3 Estimated allowance

Table A5.4 shows our assessment of YWS's efficient base expenditure for AMP8 in residential retail, as well as a comparison to Ofwat's DD models.

Table A5.4 YWS's efficient cost predictions in residential retail

	Submitted costs	Efficient modelled costs (2022–23 prices)	Efficient modelled costs (nominal)
PR24 DD models with APR update	446.0	437.9	494.1
Our proposed models with APR update	446.0	459.7	518.7
PR24 DD models without COVID-19 dummies	446.0	455.4	513.8
PR24 DD models with APR update and without COVID-19 dummies	446.0	462.1	521.5

Note: The estimated values above include a UQ benchmark, FS/RPE assumptions, as per Ofwat's DD decision.
Source: Oxera analysis.

As demonstrated in the table above, our preferred augmented model provides a material improvement of c. £21.8m (in 2022/23 prices) compared with Ofwat's PR24 DD models. The two other proposed models also show a material improvement compared with Ofwat's PR24 DD models, with values comparable to our preferred approach. The same holds true for Ofwat's PR24 DD models without COVID-19 dummies, showing a material improvement of c. £24.2m (in 2022–23 prices).

The increase in efficient modelled costs is attributed to a less stringent UQ benchmark compared with Ofwat's PR24 DD models (by approximately three percentage points on average) and the exclusion of COVID-19 dummies, which when included remove the impact of high-cost years when determining companies' forward-looking cost allowances.

A6 Regression outputs

This section presents the regression outputs for our proposed models and Ofwat's DD models following the APR24 update across all price controls.

A6.1 Wholesale water

The table below presents our proposed models for WRP.

Table A6.1 Proposed models for WRP

	WRP1	WRP2	WRP3	WRP4	WRP5	WRP6
Connected properties (log)	1.102***	1.093***	1.078***	1.073***	1.047***	1.042***
Length of mains (log)						
Water treated at complexity levels 3 to 6 (%)	0.00527***		0.00475***		0.00538***	
Weighted average treatment complexity (levels)		0.149		0.144		0.156*

	WRP1	WRP2	WRP3	WRP4	WRP5	WRP6
Booster pumping stations per length of mains (log)						
Average pumping head (log)						
Weighted average density—LAD from MSOA (log)	-1.743***	-1.545**				
Weighted average density—LAD from MSOA (log) squared	0.111***	0.0954**				
Weighted average density—MSOA (log)			-5.596***	-5.270**		
Weighted average density—MSOA (log) squared			0.342***	0.319**		
Properties per length of mains (log)					-8.349**	-7.638**
Properties per length of mains (log) squared					0.923**	0.831**
Constant	-4.839***	-5.607***	11.62	10.27	7.958	6.404

	WRP1	WRP2	WRP3	WRP4	WRP5	WRP6
Observations	221	221	221	221	221	221
RESET	0.599	0.446	0.890	0.722	0.484	0.253
BP	0	0	0	0	0	0
R-squared	0.911	0.907	0.902	0.900	0.912	0.910

Note: Significance levels: ***p<0.01, **p<0.05, *p<0.1.
Source: Oxera analysis.

The table below presents our proposed models for TWD.

Table A6.2 Proposed models for TWD

	TWD1	TWD2	TWD3	TWD4	TWD5	TWD6
Connected properties (log)				1.095***	1.059***	1.063***
Length of mains (log)	1.070***	1.018***	1.063***			
Water treated at complexity levels 3 to 6 (%)						
Weighted average treatment complexity (levels)						

	TWD1	TWD2	TWD3	TWD4	TWD5	TWD6
Booster pumping stations per length of mains (log)	0.275**	0.281**	0.365***	0.444**	0.492***	0.365***
Average pumping head (log)	0.203**	0.279***	0.205**	0.219**	0.239***	0.205**
Weighted average density—LAD from MSOA (log)	-3.059***			-2.690***		
Weighted average density—LAD from MSOA (log) squared	0.241***			0.196***		
Weighted average density—MSOA (log)		-6.410***			-6.549***	
Weighted average density—MSOA (log) squared		0.446***			0.428***	
Properties per length of mains (log)			-15.88***			-16.94***
Properties per length of mains (log) squared			2.008***			2.008***
Constant	3.925***	17.44***	26.00***	-0.694	15.79***	26.00***
Observations	221	221	221	221	221	221
RESET	0.207	0.728	0.752	0.334	0.620	0.752
BP	0	0	0	0	0	0
R-squared	0.960	0.964	0.966	0.962	0.961	0.966

Note: Significance levels: ***p<0.01, **p<0.05, *p<0.1.
Source: Oxera analysis.

The table below presents the regression results for our proposed models in WW.

Table A6.3 Proposed models for WW

	WW1	WW2	WW3	WW4	WW5	WW6
Connected properties (log)	1.085***	1.074***	1.059***	1.052***	1.048***	1.041***
Water treated at complexity levels 3 to 6 (%)	0.00276***		0.00231**		0.00289***	
Weighted average treatment complexity (levels)		0.103**		0.0950*		0.107**
Booster pumping stations per length of mains (log)	0.299*	0.317**	0.375**	0.384**	0.271*	0.272*
Average pumping head (log)	0.188*	0.169	0.196*	0.176*	0.140	0.122
Weighted average density—LAD from MSOA (log)	-2.236***	-1.987***				
Weighted average density—LAD from MSOA (log) squared	0.157***	0.139***				
Weighted average density—MSOA (log)			-5.850***	-5.274***		
Weighted average density—MSOA (log) squared			0.374***	0.336***		

	WW1	WW2	WW3	WW4	WW5	WW6
Properties per length of mains (log)					-12.36***	-11.28***
Properties per length of mains (log) squared					1.444***	1.311***
Constant	-1.952	-2.734**	13.66***	11.45***	17.14***	14.90***
Observations	221	221	221	221	221	221
RESET	0.498	0.352	0.665	0.505	0.527	0.257
BP	0	0	0	0	0	0
R-squared	0.965	0.967	0.964	0.966	0.965	0.967

Note: Significance levels: ***p<0.01, **p<0.05, *p<0.1.
Source: Oxera analysis.

The table below presents the regression results of Ofwat's DD models following the APR update for WRP.

Table A6.4 Ofwat draft determination models for WRP

	WRP1	WRP2	WRP3	WRP4	WRP5	WRP6
Connected properties (log)	1.102***	1.097***	1.078***	1.075***	1.047***	1.044***

	WRP1	WRP2	WRP3	WRP4	WRP5	WRP6
Water treated at complexity levels 3 to 6 (%)	0.00527***		0.00475***		0.00538***	
Weighted average treatment complexity (log)		0.473		0.453		0.501*
Weighted average density—LAD from MSOA (log)	-1.743***	-1.641***				
Weighted average density—LAD from MSOA (log) squared	0.111***	0.103**				
Weighted average density—MSOA (log)			-5.596***	-5.541**		
Weighted average density—MSOA (log) squared			0.342***	0.337**		
Properties per length of mains (log)					-8.349**	-8.053**
Properties per length of mains (log) squared					0.923**	0.884**
Constant	-4.839***	-5.375***	11.62	11.22	7.958	7.151
Observations	221	221	221	221	221	221
RESET	0.599	0.458	0.890	0.733	0.484	0.276
BP	0	0	0	0	0	0
R-squared	0.911	0.906	0.902	0.899	0.912	0.909

Note: Significance levels: ***p<0.01, **p<0.05, *p<0.1.

Source: Oxera analysis.

The table below presents the regression results of Ofwat's DD models following the APR update for TWD.

Table A6.5 Ofwat draft determination models for TWD

	TWD1	TWD2	TWD3	TWD4	TWD5	TWD6
Length of mains (log)	1.074***	1.026***	1.071***	1.070***	1.017***	1.049***
Booster pumping stations per length of mains (log)	0.352**	0.366***	0.438***			
Average pumping head (log)				0.242***	0.311***	0.238***
Weighted average density—LAD from MSOA (log)	-2.907***			-3.297***		
Weighted average density—LAD from MSOA (log) squared	0.230***			0.253***		
Weighted average density—MSOA (log)		-5.967***			-7.221***	
Weighted average density—MSOA (log) squared		0.419***			0.491***	
Properties per length of mains (log)			-15.59***			-18.22***

	TWD1	TWD2	TWD3	TWD4	TWD5	TWD6
Properties per length of mains (log) squared			1.978***			2.250***
Constant	4.523***	17.13***	26.42***	3.660**	19.72***	30.05***
Observations	221	221	221	221	221	221
RESET	0.0996	0.210	0.680	0.249	0.685	0.822
BP	0	0	0	0	0	0
R-squared	0.955	0.955	0.960	0.956	0.960	0.961

Note: Significance levels: ***p<0.01, **p<0.05, *p<0.1.
Source: Oxera analysis.

The following tables present the regression results of Ofwat's DD models following the APR update for WW.

Table A6.6 Ofwat draft determination models for WW (WW1–WW6)

	WW1	WW2	WW3	WW4	WW5	WW6
Connected properties (log)	1.085***	1.075***	1.063***	1.057***	1.052***	1.046***
Water treated at complexity levels 3 to 6 (%)	0.00309***		0.00265**		0.00318***	

	WW1	WW2	WW3	WW4	WW5	WW6
Weighted average treatment complexity (log)		0.367**		0.339**		0.381**
Booster pumping stations per length of mains (log)	0.378**	0.391***	0.439***	0.443***	0.316**	0.312**
Average pumping head (log)						
Weighted average density—LAD from MSOA (log)	-2.017***	-1.825***				
Weighted average density—LAD from MSOA (log) squared	0.142***	0.128***				
Weighted average density—MSOA (log)			-5.273***	-4.875***		
Weighted average density—MSOA (log) squared			0.338***	0.311***		
Properties per length of mains (log)					-11.94***	-11.15***
Properties per length of mains (log) squared					1.394***	1.295***
Constant	-1.627	-2.387	12.44***	10.75**	17.00***	15.17***
Observations	221	221	221	221	221	221
RESET	0.246	0.114	0.379	0.211	0.306	0.136
BP	0	0	0	0	0	0
R-squared	0.964	0.966	0.963	0.965	0.965	0.966

Note: Significance levels: ***p<0.01, **p<0.05, *p<0.1.
Source: Oxera analysis.

Table A6.7 Ofwat draft determination models for WW (WW7–WW12)

	WW7	WW8	WW9	WW10	WW11	WW12
Connected properties (log)	1.087***	1.078***	1.056***	1.049***	1.038***	1.032***
Water treated at complexity levels 3 to 6 (%)	0.00242***		0.00192*		0.00265***	
Weighted average treatment complexity (log)		0.311*		0.288		0.339**
Booster pumping stations per length of mains (log)						
Average pumping head (log)	0.240**	0.242**	0.248**	0.250**	0.177	0.175
Weighted average density—LAD from MSOA (log)	-2.564***	-2.424***				
Weighted average density—LAD from MSOA (log) squared	0.176***	0.166***				
Weighted average density—MSOA (log)			-7.150***	-6.839***		
Weighted average density—MSOA (log) squared			0.450***	0.429***		
Properties per length of mains (log)					-14.18***	-13.53***
Properties per length of mains (log) squared					1.635***	1.555***
Constant	-2.046	-2.675	17.50***	16.14***	20.29***	18.79***
Observations	221	221	221	221	221	221
RESET	0.598	0.567	0.715	0.762	0.660	0.524
BP	0	0	0	0	0	0

	WW7	WW8	WW9	WW10	WW11	WW12
R-squared	0.959	0.960	0.955	0.956	0.961	0.963

Note: Significance levels: ***p<0.01, **p<0.05, *p<0.1.
Source: Oxera analysis.

A6.2 Wholesale WWNP

The table below presents the regression results for our proposed models in wholesale WWNP.

Table A6.8 Proposed models for WWNP

	SWC1	SWC2	SWC3	SWT1	WWNP1
Sewer length (log)	0.786***	0.872***	0.834***		
Pumping capacity per length (log)	0.347**	0.615***	0.557***		0.332***
Properties per length (log)	1.306***				
Urban MSOA rainfall per length (log)	0.0862***	0.141***	0.140***		0.0739**
Weighted average density—LAD from MSOA (log)		0.295***			
Weighted average density—MSOA (log)			0.496***		

	SWC1	SWC2	SWC3	SWT1	WWNP1
Sewage load (log)				0.814***	0.723***
Weighted average sewage treatment size (log)				-0.238***	-0.0898***
Proportion of ammonia permit <3mg/l (%)				0.00581***	0.00576***
Constant	-8.304***	-6.440***	-7.799***	-3.167***	-2.680***
Observations	130	130	130	130	130
RESET	0.0222	0.0205	0.0185	0.888	0.00860
BP	1.46e-06	0	0	0	0.000515
R-squared	0.910	0.901	0.897	0.895	0.947

Note: Significance levels: ***p<0.01, **p<0.05, *p<0.1.
Source: Oxera analysis.

The table below presents the regression results of Ofwat's DD models following the APR update for WWNP.

Table A6.9 Ofwat draft determination models for WWNP

	SWC1	SWC2	SWC3	SWT1	SWT2	WWNP1	WWNP2
Sewer length (log)	0.786***	0.872***	0.834***				
Pumping capacity per length (log)	0.347**	0.615***	0.557***			0.423***	0.332***
Properties per length (log)	1.306***						
Urban MSOA rainfall per length (log)	0.0862***	0.141***	0.140***			0.0681**	0.0739**
Weighted average density—LAD from MSOA (log)		0.295***					
Weighted average density—MSOA (log)			0.496***				
Sewage load (log)				0.709***	0.814***	0.751***	0.723***
Weighted average sewage treatment size (log)					-0.238***		-0.0898**
Percentage of load treated in bands 1-3 (%)				0.0283		0.0240***	
Proportion of ammonia permit <3mg/l (%)				0.00547***	0.00581***	0.00523***	0.00576***
Constant	-8.304***	-6.440***	-7.799***	-4.227***	-3.167***	-4.042***	-2.680***
Observations	130	130	130	130	130	130	130
RESET	0.0222	0.0205	0.0185	0.159	0.888	0.158	0.00860
BP	1.46e-06	0	0	0	0	4.33e-06	0.000515

	SWC1	SWC2	SWC3	SWT1	SWT2	WWNP1	WWNP2
R-squared	0.910	0.901	0.897	0.843	0.895	0.944	0.947

Note: Significance levels: ***p<0.01, **p<0.05, *p<0.1.
Source: Oxera analysis.

A6.3 Bioresources

The table below presents the regression results of Ofwat's DD models following the APR update for bioresources.

Table A6.10 Ofwat draft determination models for bioresources

	BR1	BR2	BR3	BR4
Percentage of load treated in bands 1-3 (%)	0.0519***			
Weighted average density—LAD from MSOA (log)		-0.241*		
Weighted average density—MSOA (log)			-0.340*	
Number of SWTs per property (log)				0.203**
Constant	-0.863***	1.073	2.016	0.997
Observations	130	130	130	130
RESET	0.705	0.00242	0.0460	0.435

	BR1	BR2	BR3	BR4
BP	0	0	0	0
R-squared	0.243	0.151	0.124	0.143

Note: Significance levels: ***p<0.01, **p<0.05, *p<0.1.
Source: Oxera analysis.

A6.4 Residential retail

The table below presents the regression results of Ofwat's DD models following the APR update for residential retail.

Table A6.11 Ofwat draft determination models for residential retail

	RDC1	RDC2	ROC1	ROC2	RTC1	RTC2	RTC3	RTC4
Average bill size (£ per/household) (log)	1.065***	1.013***			0.660***	0.692***	0.549***	0.566***
Percentage of households with payment default (%)	0.0419***				0.0226***		0.0235***	
Income deprivation score (interpolated) (%)		0.0670***				0.0279**		0.0317**

	RDC1	RDC2	ROC1	ROC2	RTC1	RTC2	RTC3	RTC4
Proportion of dual-service households (%)			0.00191**	0.00396***				
Total number of households (log)				-0.0885**	-0.0928***	-0.0960***		
COVID-19 dummy for 2019–20 (nr)	0.369***	0.388***			0.176***	0.180***	0.178***	0.184***
COVID-19 dummy for 2020–21 (nr)	0.191**	0.201***			0.0574*	0.0574**	0.0568*	0.0586**
Constant	-4.706***	-4.289***	2.881***	4.025***	0.426	0.462	-0.254	-0.203
Observations	187	187	187	187	187	187	187	187
Adjusted R-squared	0.667	0.667	0.109	0.117	0.708	0.645	0.666	0.640
RESET	0.281	0.285	0.931	0.132	0.701	0.429	0.493	0.305
BP	0	0	0	0	0	0	0	0

Note: Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$
Source: Oxera analysis.

The table below presents the regression results of Ofwat's DD models without COVID-19 dummies and following the APR update for residential retail.

Table A6.12 Ofwat draft determination models for residential retail without COVID-19 dummies

	RDC1	RDC2	ROC1	ROC2	RTC1	RTC2	RTC3	RTC4
Average bill size (£ per/household) (log)	1.117***	1.112***			0.687***	0.772***	0.575***	0.628***
Percentage of households with payment default (%)	0.0308**				0.0180**		0.0177**	
Income deprivation score (interpolated) (%)		0.0346				0.00616		0.00993
Proportion of dual-service households (%)			0.00191**	0.00396***				
Total number of households (log)				-0.0885**	-0.0930***	-0.100***		
Constant	-4.693***	-4.384***	2.881***	4.025***	0.404	0.370	-0.245	-0.255
Observations	187	187	187	187	187	187	187	187
Adjusted R-squared	0.633	0.628	0.109	0.117	0.669	0.606	0.638	0.615
RESET	0.117	0.0763	0.931	0.132	0.00880	0.0738	0.000762	0.000134
BP	0	0	0	0	0	0	0	0

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1
Source: Oxera analysis.

The table below presents the regression results when we smooth doubtful debt costs over the modelling period, in line with Ofwat's treatment of depreciation, and following the APR update for residential retail.

Table A6.13 Proposed models for residential retail (Approach 1)

	RDC1	RDC2	ROC1	ROC2	RTC1	RTC2	RTC3	RTC4
Average bill size (£ per/household) (log)	0.855***	0.886***			0.639***	0.708***	0.542***	0.582***
Percentage of households with payment default (%)	0.0382***				0.0211***		0.0230***	
Income deprivation score (interpolated) (%)		0.0715***				0.0249**		0.0291**
Proportion of dual-service households (%)			0.00191**	0.00396***				
Total number of households (log)				-0.0885**	-0.0893***	-0.105***		
Constant	-3.346***	-3.540***	2.881***	4.025***	0.552*	0.556	-0.181	-0.236
Observations	187	187	187	187	187	187	187	187
Adjusted R-squared	0.783	0.784	0.109	0.117	0.735	0.665	0.692	0.670

	RDC1	RDC2	ROC1	ROC2	RTC1	RTC2	RTC3	RTC4
RESET	0.0563	0.0151	0.931	0.132	0.220	0.863	0.110	0.148
BP	0	0	0	0	0	0	0	0

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1
Source: Oxera analysis.

The table below presents the regression results when use linear interpolation to impute the doubtful debt costs in 2020 and 2021, and following the APR update for residential retail.

Table A6.14 Proposed models for residential retail (Approach 2)

	RDC1	RDC2	ROC1	ROC2	RTC1	RTC2	RTC3	RTC4
Average bill size (£ per/household) (log)	1.123***	1.042***			0.678***	0.717***	0.566***	0.587***
Percentage of households with payment default (%)	0.0467**				0.0222***		0.0229***	
Income deprivation score (interpolated) (%)		0.0826**				0.0263*		0.0295*

	RDC1	RDC2	ROC1	ROC2	RTC1	RTC2	RTC3	RTC4
Proportion of dual-service households (%)			0.00191**	0.00396***				
Total number of households (log)				-0.0885**	-0.0933***	-0.0977***		
Constant	-5.171***	-4.672***	2.881***	4.025***	0.336	0.364	-0.340	-0.298
Observations	187	187	187	187	187	187	187	187
Adjusted R-squared	0.639	0.647	0.109	0.117	0.718	0.654	0.680	0.654
RESET	0.347	0.127	0.931	0.132	0.000141	0.00120	0.000320	1.15e-05
BP	0	0	0	0	0	0	0	0

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1
Source: Oxera analysis

The table below presents the regression results when we impute the doubtful debt costs in 2020 and 2021 using the average doubtful debt costs in 2017–19, and following the APR update for residential retail.

Table A6.15 Proposed models for residential retail (Approach 3)

	RDC1	RDC2	ROC1	ROC2	RTC1	RTC2	RTC3	RTC4
Average bill size (£ per/household) (log)	1.069***	1.013***			0.671***	0.715***	0.563***	0.588***
Percentage of households with payment default (%)	0.0470***				0.0223***		0.0233***	
Income deprivation score (interpolated) (%)		0.0796***				0.0266*		0.0302**
Proportion of dual-service households (%)			0.00191**	0.00396***				
Total number of households (log)				-0.0885**	-0.0933***	-0.0977***		
Constant	-5.171***	-4.672***	2.881***	4.025***	0.336	0.364	-0.340	-0.298
Observations	187	187	187	187	187	187	187	187
Adjusted R-squared	0.639	0.647	0.109	0.117	0.718	0.654	0.680	0.654
RESET	0.347	0.127	0.931	0.132	0.000141	0.00120	0.000320	1.15e-05
BP	0	0	0	0	0	0	0	0

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1
Source: Oxera analysis.

Contact

Dr Srini Parthasarathy
Partner
+44 (0) 20 7776 6612
srini.parthasarathy@oxera.com

oxera.com