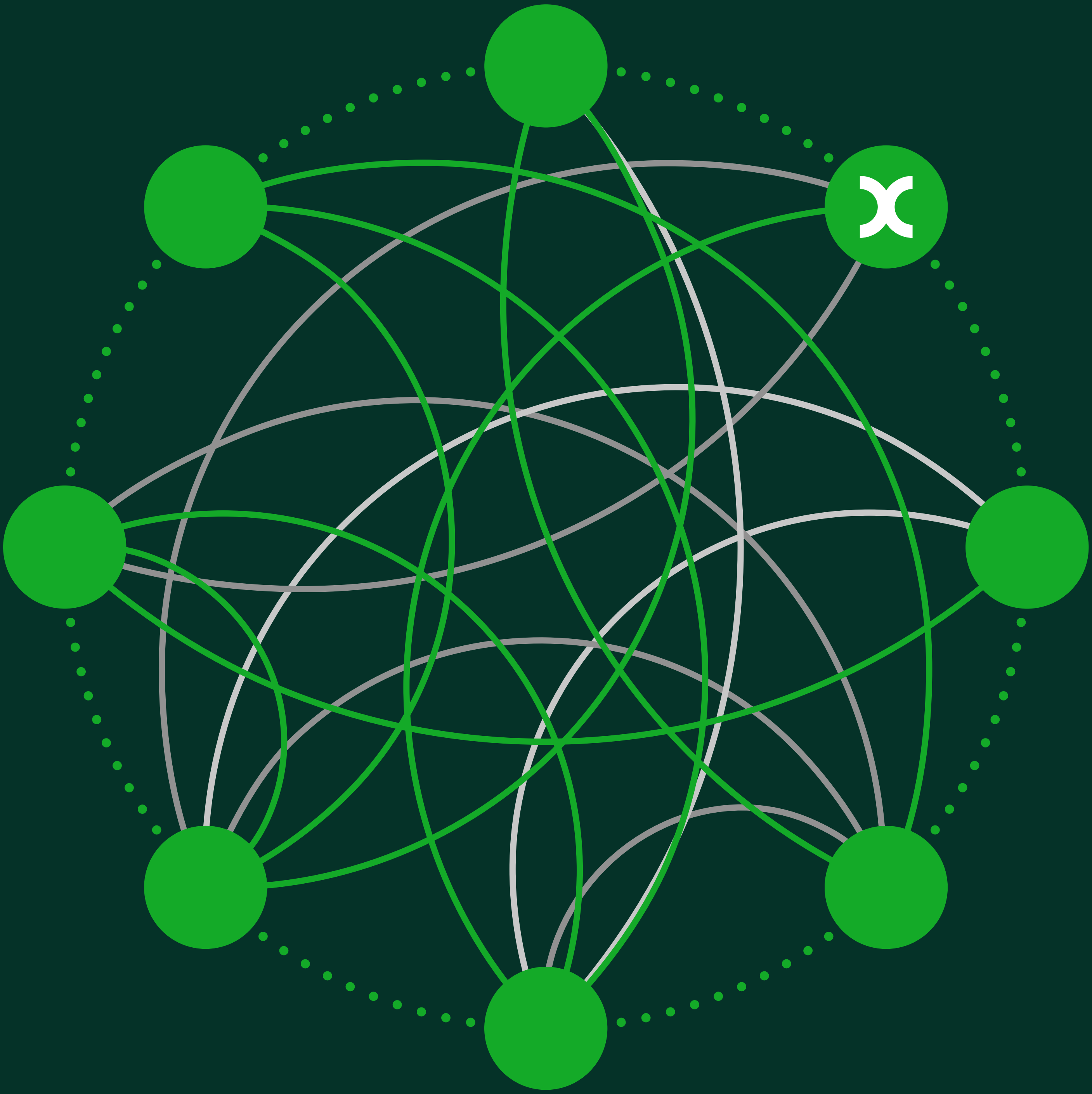
Appendix
YKY21_Oxera cost modelling

An assessment of Yorkshire Water Services' base cost requirements

—

Prepared for Yorkshire Water

28 September 2023



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; in Belgium, no. 0651 990 151, branch office: Avenue Louise 81, 1050 Brussels, Belgium; and in Italy, REA no. RM - 1530473, branch office: Via delle Quattro Fontane 15, 00184 Rome, Italy. Oxera Consulting (France) LLP, a French branch, registered office: 60 Avenue Charles de Gaulle, CS 60016, 92573 Neuilly-sur-Seine, France and registered in Nanterre, RCS no. 844 900 407 00025. Oxera Consulting (Netherlands) LLP, a Dutch branch, registered office: Strawinskylaan 3051, 1077 ZX Amsterdam, The Netherlands and registered in Amsterdam, KvK no. 72446218. Oxera Consulting GmbH is registered in Germany, no. HRB 148781 B (Local Court of Charlottenburg), registered office: Rahel-Hirsch-Straße 10, Berlin 10557, Germany.

Figures and Tables

Executive summary

Yorkshire Water (YWS) has commissioned Oxera to provide a robust top-down assessment of its efficient base expenditure requirements in AMP8 (2026–30). This evaluation relies on two primary sources of information.

First, we have developed a suite of econometric cost models that builds on the PR24 base modelling consultation models presented by Ofwat in April 2023,[1] incorporating the latest data from the 2023 Annual Performance Reports (APRs)[2], feedback from the industry on Ofwat's consultation models,[3] and our own empirical investigation of the augmented dataset. As with Ofwat's consultation models, these augmented models seek to capture a wide range of industry-wide characteristics, as well as cost pressures specific to YWS, and are aligned with Ofwat's modelling criteria.

Second, we have used the PR24 consultation models presented by Ofwat on the augmented dataset to derive an efficient baseline cost prediction. As these models do not capture some of the specific cost pressures that YWS faces or is expected to face over AMP8 adequately, we have supplemented the baseline prediction from the PR24 consultation models with cost adjustments to ensure that YWS is able to recover the efficient cost of its operations. The specific 'cost adjustment claims' (CACs) that we have explored for YWS are as follows.

- **Combined sewers.** YWS has the second-highest proportion of combined sewers in the industry. Combined sewers are associated with higher costs due to the additional risk of flooding associated with combined sewers versus separate sewers. Moreover, the prevalence of combined sewers on the sewerage network is driven largely by the prevalence of combined sewers at the time of privatisation—i.e. their presence relates largely to legacy decisions and, as such, the degree or risk of endogeneity is very low. As combined sewers are not explicitly captured in the PR24 consultation models, the PR24 consultation models will underfund YWS on this basis.
- **Phosphorus-removal.** YWS is anticipating a material increase in phosphorus-removal (P-removal) activity in AMP8 due to legislative requirements. As noted by Ofwat in the PR24 modelling consultation,

---

[1] Ofwat (2023), 'Econometric base cost models for PR24', April.
[2] YWS shared a dataset containing relevant APR data that was compiled by the industry.
[3] See Ofwat (2023), 'Econometric base cost modelling responses', available at: https://www.ofwat.gov.uk/consultation/pr24-econometric-base-cost-models-consultation/#Responses.

P-removal activity is not adequately captured in the consultation models.[4] Therefore, we have built on a proposal from Ofwat to quantify an adjustment to YWS's baseline efficient expenditure to reflect anticipated P-removal activity.

Note that these CACs relate to limitations with Ofwat's PR24 consultation models. In our augmented models, we include cost drivers that can account for these characteristics explicitly and, as such, additional CACs may not be required if such models are adopted. The details surrounding these CACs—including their justification and quantification—are outlined in a separate report.[5]

As part of our econometric modelling exercise, we have also undertaken an initial investigation of the reliability and uncertainty associated with the models and data using objective, scientific methods.[6] Following the principle outlined by the Competition and Markets Authority (CMA) in the PR19 redetermination—namely, that the benchmark should be informed by the quality of the econometric models[7] (and, therefore, their inherent limitations)—we have used the scientific methods outlined below to inform the choice of the benchmark.[8]

First, we have examined the width of the confidence intervals around companies' cost predictions, which is a direct measure of the level of uncertainty in the models.[9] This technique has been considered by regulators to inform the benchmark and was also investigated by the CMA in the PR19 redetermination.[10]

While this approach has precedent and provides a measure of uncertainty based on Ofwat's own modelling assumptions, it does not in itself provide an assessment of exactly what the appropriate benchmark should be; rather, it

---

[4] Ofwat (2023), 'Econometric base cost models for PR24', April, section 4.

[5] See Oxera (2023), 'An assessment of Yorkshire Water's cost adjustment claims', September.

[6] In addition to the methods outlined in this report, Monte Carlo analysis could be used to assess the impact of data uncertainty on the models and companies' performance. This was explored as part of the PR19 redetermination, and the CMA concluded that such an approach has merit in principle. However, the CMA argued that it required a value judgement regarding the level of noise in the data (see CMA (2020), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report', March, para. 4.399).

[7] For example, see CMA (2021), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report', March, para. 4.493.

[8] Determining the benchmark based on the quality of the models is a well-established method in regulatory applications and rulings from the CMA. For example, in the redetermination of Bristol Water's allowance at PR14, the CMA applied an average benchmark to account for the issues that it had identified with both Ofwat's PR14 models and the models that the CMA developed for the inquiry.

[9] In particular, the wider (narrower) the confidence interval, the more (less) uncertainty there is in the models. An assessment of this noise (uncertainty) to signal (inefficiency) ratio in comparison with past regulatory decisions in water and other sectors in the UK and elsewhere can be used to inform the level of the benchmark or an acceptable correction for uncertainty.

[10] See CMA (2020), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report', March, pp. 202–210.

can assess whether the uncertainty is higher or lower than the uncertainty in previous regulatory decisions, which could support whether the benchmark should be less or more stringent. That is, the method requires an anchor based on past regulatory decisions. As noted by Ofwat, the PR24 consultation models build on the PR19 models[11] and, as such, it is appropriate to compare the confidence intervals in the models outlined in this report with the confidence intervals in the PR19 models.

Second, we have employed stochastic frontier analysis (SFA) to assess the level of statistical noise (e.g. data and modelling errors) in the models. As this technique provides a data-driven assessment of the amount of noise in the models, it does not require the same reliance on previous regulatory decisions as the confidence interval analysis. SFA is one of the most commonly used econometric methods for efficiency assessment by regulators across Europe[12] and has also been considered by UK regulators to assess the level of uncertainty in models.[13] It was also investigated by the CMA as part of the PR19 redetermination. While the results from SFA were consistent with the results from confidence interval analysis on the PR19 models, the CMA observed that the conclusions from SFA were sensitive to some of the assumptions made. To mitigate this issue, we have explored multiple SFA models as part of this report.

The cost models developed on the augmented data, 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.

Wholesale water

The augmented models we have developed as part of our assessment differ from the PR24 consultation models in the following respects.

- **Scale.** Ofwat used length of mains as the scale driver in all of its treated water distribution (TWD) consultation models. As Ofwat's modelled cost includes costs associated with population growth, such as network reinforcement, we control for connected properties

---

[11] Ofwat (2023), 'Econometric base cost models for PR24', April, section 1.

[12] For example, SFA has been the main econometric method used by the Bundesnetzagentur (the German energy regulator) to estimate the static efficiency of German electricity distribution system operators since the introduction of incentive regulation. See Bundesnetzagentur (2018), 'Decision BK4-18-056', November. SFA is also employed by several other European regulators across sectors (alongside methods such as data envelopment analysis and variants of ordinary least squares analysis).

[13] For example, ORR has used SFA to assess the efficiency of both Network Rail and Highways England. See ORR (2013), 'PR13 Efficiency Benchmarking of Network Rail using LICB', August; and ORR (2017), 'Benchmarking regional maintenance costs on England's Strategic Road Network').

as a scale variable in half of our TWD models,[14] with the remaining half controlling for length of mains.

- **Population density.** In its water resources plus (WRP), TWD and WW consultation models, Ofwat controlled for three different measures of density: (i) properties per length of main; (ii) MSOA-based weighted average density;[15] and (iii) LAD-based weighted average density.[16] We do not consider that properties per length of main is a relevant driver of WRP costs, given that this driver relates entirely to TWD assets. Moreover, we consider at this stage that the MSOA-based weighted average density measure is preferable to the LAD-based measure, given that the former is a more granular version of the latter.[17]
- **Treatment complexity.** Ofwat's weighted average complexity (WAC) is a weighted proportion variable and, as such, should be modelled in levels rather than in logarithms (Ofwat currently does the latter). Doing so is consistent with Ofwat's approach to modelling proportion variables, improves the interpretability of the estimated coefficient, and leads to an improvement in the statistical quality of the model. Therefore, we control for WAC in levels rather than in logarithms.
- **Topography.** In its cost assessment consultation models, Ofwat controls for *either* average pumping head (APH) *or* booster pumping stations per length of mains. On the current dataset, both drivers could potentially be included in the models jointly as the two drivers could capture different aspects of operational costs. While this proposition would benefit from validation from an operational perspective, we note that including both drivers results in improved statistical performance of the models. Ofwat could consider this as an alternative to considering them separately and triangulating the results. As such, both booster pumping stations per length of mains and APH are included in our augmented TWD and WW models. Any possible bias stemming from joint or separate modelling of the two drivers for any company can be addressed through the CAC framework.

In both the PR24 consultation models and our augmented models, the confidence interval analysis shows that the models predict companies' expenditure with a level of uncertainty comparable to the models used by the

---

[14] We find that network reinforcement expenditure is correlated more strongly with connected properties than with length of mains.

[15] Middle layer Super Output Area (MSOA).

[16] Local Authority Districts (LAD).

[17] The optimal granularity of the density driver will depend on exactly how density is related to costs from an operational perspective. That is, the LAD-based weighted average density measure may be operationally superior to the MSOA-based measure if the costs associated with density or sparsity are expected to manifest at a lower level of granularity. Given that the two measures perform similarly in the models from a statistical perspective, the selection of the weighted average density measure may require additional operational evidence.

CMA in the PR19 redetermination, where an upper-quartile (UQ) benchmark was applied. Specifically, the WRP and WW models predict companies' costs with a higher degree of uncertainty, while the TWD models predict companies' costs with a lower degree of uncertainty than the CMA PR19 models.

Moreover, the SFA modelling under various assumptions suggests that the models are estimated with significant uncertainty, indicating that even a UQ benchmark may be inconsistent with the evidence regarding the level of noise and uncertainty in the modelling. Therefore, throughout this report, we consider three benchmarks: (i) the UQ (75th percentile), in line with precedent from PR19; (ii) the upper tercile (i.e. the upper third or 66th percentile); and (iii) the average benchmark (50th percentile).

The table below shows our assessment ('Augmented models') of YWS's efficient base expenditure for AMP8 (2026–30), as well as a comparison with the PR24 consultation models ('PR24 models').

Assessment of YWS's efficient AMP8 base expenditure—wholesale water

	Augmented models	**PR24 models**
Average efficiency	£1,763m	£1,744m
Upper-quartile	£1,762m	£1,730m
Upper-tercile	£1,775m	£1,764m

Note: Expenditure is expressed in 2022/23 prices.
Source: Oxera analysis.

The table shows that YWS's efficient expenditure is c. £1,762m–£1,775m, depending on the benchmark applied. Note that this is slightly higher than the range under the PR24 consultation models, indicating that the PR24 consultation models may underestimate YWS's efficient cost allowance.

Wholesale wastewater (network plus)

The augmented models that we have developed for YWS differ from the PR24 consultation models in the following respects.

- **Population density.** As per the WW models, we consider that properties per length of sewer and MSOA-based weighted average density are more appropriate density drivers of sewage collection (SWC) costs. We do not include the LAD-based weighted average density measure in our augmented models, given that the MSOA-based measure is more granular.

- **Treatment complexity.** As observed above, Ofwat's consultation models do not account for P-removal activity. Therefore, we control for a composite complexity variable, defined as the weighted sum of P-removal activity and ammonia-removal (N-removal) activity in all of our sewage treatment (SWT) and WWNP models. In this way, our augmented models can better reflect the increasing cost pressure that YWS faces with respect to increased P-removal activity.
- **Combined sewers.** We control for the proportion of combined sewers as a network complexity driver in our SWC and WWNP models. These are legacy assets, inherited by companies at privatisation, and are associated with higher costs due to the additional risk of flooding. That is, the driver is exogenous in the short term and medium term and is a relevant driver of costs from an operational perspective. This driver performs well in the econometric models (it is statistically significant and improves model fit relative to Ofwat's PR24 consultation models) and can capture (to some extent) the increased costs that YWS faces as a result of combined sewers.
- **Economies of scale.** Ofwat controls for three measures of sewage treatment works (STW)-level economies of scale: (i) load treated in size bands 1 to 3 (%); (ii) load treated in STWs ≥ 100,000 people (%); and (iii) weighted average treatment size (WATS). We control for WATS in all of our SWT and WWNP models, given that it performs better in the models from a statistical perspective on the current dataset and is less reliant on arbitrary thresholds that the two other measures require.
- **Urban rainfall.** Ofwat controls for urban rainfall in half of its relevant consultation models to account for the increased costs that companies face as a result of weather conditions. Given that urban rainfall is an operationally relevant driver of expenditure that performs well in the models, we control for urban rainfall in all of our augmented models.[18]

The confidence interval analysis shows that the sewage collection (SWC) and sewage treatment (SWT) models predict companies' costs with a higher degree of uncertainty than the equivalent models at PR19. Meanwhile, the WWNP models predict companies' costs with a lower degree of uncertainty

---

[18] The urban rainfall measure developed by Ofwat is defined as the total rainfall in a company's operating region multiplied by the proportion of a company's operating region that is urban. Alternative measures of urban rainfall (e.g. ones that account for the location of rainfall, or ones that account for 'peak' rainfall or 'severe weather') may be more appropriate and could be developed ahead of PR24. Nonetheless, at this stage, we consider that the biases associated with controlling for an imperfect measure of urban rainfall may be lower than the biases associated with omitting urban rainfall entirely.

than the SWC and SWT models; however, robust comparisons cannot be made with PR19 given that the CMA did not use WWNP models at PR19.

Some SFA models indicate that the average efficiency gap in some SFA models is similar to that under a UQ benchmark. However, other SFA models indicate that a UQ benchmark is overly stringent.

Given the mixed evidence,[19] we consider at this stage that somewhere between a UQ and an upper-tercile benchmark might be appropriate. The table below shows YWS's estimated efficient cost allowance under the three benchmarks considered in this report.

Assessment of YWS's efficient base expenditure—wastewater (network plus)

	Augmented models	PR24 models
Average efficiency	£2,170m	£1,774m
Upper-quartile	£2,127m	£1,765m
Upper-tercile	£2,147m	£1,792m

Note: Expenditure is expressed in 2022/23 prices.
Source: Oxera analysis.

The table shows that, on the basis of the augmented models, YWS's efficient expenditure is c. £2,127m–£2,170m, depending on the benchmark applied. This is materially higher than the range indicated by Ofwat's consultation models; however, this is expected as the consultation models do not adequately capture YWS-specific cost pressures (namely, combined sewers and P-removal activity), some of which Ofwat had also acknowledged in the modelling consultation.[20]

The table below shows YWS's efficient cost allowance once these CACs are added to the PR24 consultation models' predictions. While the determination of the appropriate benchmark will require further work, in line with regulatory precedent and the evidence presented above, we have currently applied a UQ

---

[19] Specifically: (i) the confidence interval analysis shows that a less stringent benchmark may be required in the SWC and SWT models (relative to PR19); (ii) the confidence interval analysis suggests that a relatively more stringent benchmark may be appropriate in the WWNP models; (iii) some SFA models broadly support an upper-quartile benchmark; and (iv) alternative SFA models indicate that a less stringent benchmark would be appropriate.

[20] Ofwat (2023), 'Econometric base cost models for PR24', April, section 4.

benchmark to the relevant CACs[21] to ensure that they are efficient and can be added to the efficient baseline costs, which is also corrected for the UQ efficiency challenge for consistency.

Assessment of YWS's efficient base expenditure—wastewater (network plus)

	Augmented models	PR24 models
WWNP (UQ)	£2,127m	£1,765m
CAC1: Combined sewers	-	£88m
CAC2: P-removal	-	£110m
Overall AMP8 cost prediction	£2,127m	£1,963m

Note: Expenditure is expressed in 2022/23 prices.
Source: Oxera analysis.

The table shows that the incorporation of relevant CACs narrows the gap between YWS's efficient cost prediction in the augmented models and its efficient cost prediction in the PR24 consultation models. Nonetheless, a material gap remains, indicating that the PR24 consultation models may have other biases that lower YWS's efficient cost prediction, given that the augmented models improve the economic, operational and statistical quality of the models.

Bioresources

On the augmented dataset, we have explored additional cost drivers and model specifications in Ofwat's BR consultation models, but have not identified models that are objectively superior to those presented by Ofwat at the consultation. Therefore, we have focused solely on Ofwat's PR24 consultation models when assessing YWS's efficient BR cost prediction. However, the cost driver specifications may require further development as additional outturn data and business plan data becomes available.

The confidence intervals around companies' BR cost predictions are materially wider than in the other wholesale controls, and materially wider than the equivalent models at PR19. Moreover, the SFA models indicate that all of the estimated (in)efficiency in Ofwat's consultation models could be

---

[21] Specifically, we have applied a UQ efficiency challenge to the combined sewers CAC. Meanwhile, the P-removal CAC estimate is derived from YWS's bottom-up analysis, and is supported by top-down modelling, so we do not apply an additional efficiency challenge.

driven by statistical noise. In this context, a less stringent benchmark should be applied to the BR models than the other controls. Indeed, the SFA modelling indicates that even an average benchmark may overestimate the scope for efficiency improvements.

The table below shows YWS's efficient cost prediction in the BR models under different benchmark assumptions, with the caution that these benchmarks may overestimate the scope for efficiency improvements following the evidence outlined above.

Assessment of YWS's efficient base expenditure—wastewater (bioresources)

	PR24 models
Average efficiency	£416m
Upper-quartile	£376m
Upper-tercile	£423m

Note: Expenditure is expressed in 2022/23 prices.
Source: Oxera analysis.

Residential retail

On the augmented dataset, we have explored additional cost drivers and model specifications in the residential retail models, including composite deprivation metrics, population transience and metered households. However, the inclusion of these cost drivers does not lead to an objective improvement in the quality of the PR24 consultation models on the current dataset. Therefore, we do not amend Ofwat's cost driver specifications in our augmented models. However, as per the BR models, the cost driver specifications may need to be revisited in light of new outturn and business plan data.

A key concern with Ofwat's residential retail consultation models is the use of time dummies to account for the spike in doubtful debt costs during the COVID-19 period. While the increase in doubtful debt expenditure needs to be accounted for to ensure that a sensible statistical relationship can be estimated, we consider that time dummies are blunt instruments as they capture several effects and not just the intended issue. The time dummies essentially remove the impact of high-cost years when determining companies' forward-looking cost allowances.[22] This also assumes that there will be no event like COVID-19 in AMP8 that could cause an increase in doubtful debt, and that there are no persistent effects of COVID-19 on efficient retail costs. We are unaware of any evidence to suggest that either of these assumptions is likely.

We have therefore explored alternative methods for accounting for the spike in doubtful debt that are targeted and mitigate the purported need for time dummies. In our augmented models, we smooth the costs of doubtful debt over the modelling period. Doing so leads to a material improvement in the statistical quality of the models in terms of model fit and statistical significance. This is consistent with Ofwat's (and other regulators') approaches to modelling expenditure that is subject to volatility, such as depreciation.

Our analysis shows that the confidence intervals under the augmented models are narrower than those under the PR24 consultation models, indicating that the augmented models predict companies' expenditure with a higher degree of uncertainty. Moreover, the confidence intervals in the augmented models are broadly comparable to those in the CMA's WW models in the PR19 determination, where a UQ benchmark was applied. Therefore, there is no strong evidence to support a benchmark that is more stringent than the UQ.

The SFA modelling is less conclusive—some SFA models show that a UQ benchmark may be broadly appropriate in the total retail cost (RTC) models, while a significantly less stringent benchmark is required in the retail debt collection (RDC) models. However, alternative SFA models show that a UQ benchmark is overly stringent in all residential retail models. Therefore, as per the WW models, we present YWS's allowance under three benchmarks: (i) average; (ii) upper tercile; and (iii) UQ; as shown in the table below.

---

[22] This is assuming that Ofwat sets the dummies equal to zero over the forecast period.

	Augmented models	PR24 models
Average efficiency	£494m	£491m
Upper-quartile	£467m	£447m
Upper-tercile	£469m	£462m

Note: Expenditure is expressed in 2022/23 prices.
Source: Oxera analysis.

Concluding remarks

The analysis presented in this report is anchored largely on outturn-based cost models. We note that the publication of forecast data in companies' PR24 business plans will allow Ofwat (and companies) to assess the ability of the models to predict future expenditure requirements. The ability (or inability) of the models to predict forward-looking expenditure requirements may necessitate additional data processing and developing alternative cost models. Moreover, forecast information could be incorporated directly in the cost models, in line with Ofgem's approach to cost assessment in the RIIO-2 price controls.[23] This modelling may also affect the optimal selection of cost drivers.

Our analysis also assumes that the general operating environment (including the macroeconomic environment and service obligations) will be broadly stable between the historical modelling period and AMP8. Where there are known cost pressures that can be readily modelled, we have sought to account for these directly in the cost models or through the CAC process (e.g. relating to P-removal and combined sewers). However, if there are other material changes to the operating environment in AMP8, such as more stringent service performance targets, then additional allowances may be required.

---

[23] For example, see Ofgem (2022)', 'RIIO-ED2 Final Determinations Core Methodology Document', November.

1 Introduction

Yorkshire Water (YWS) has commissioned Oxera to provide a robust, top-down assessment of its efficient base expenditure requirements in AMP8 (2026–30). As part of our assessment, we consider two sources of evidence.

First, we have developed a suite of econometric cost models that builds on the PR24 base modelling consultation models presented by Ofwat in April 2023,[24] incorporating the latest data from the 2023 Annual Performance Reports (APRs)[25], feedback from the industry on Ofwat's consultation models,[26] and our own empirical investigation on the augmented dataset. As with Ofwat's consultation models, these models seek to capture a wide range of industry-wide characteristics, as well as cost pressures specific to YWS, and are aligned with Ofwat's modelling criteria. The estimation and development of robust econometric models is the focus of this report.

Second, we have reviewed whether the efficient cost predictions derived in the first step need to be adjusted to reflect YWS-specific operating characteristics through the cost adjustment claim (CAC) process. We have examined two of YWS's proposed CACs: one relating to an increase in P-removal activity in AMP8, and the other relating to the prevalence of combined sewers on YWS's network. The details of this analysis are presented in a separate report.[27]

This report is structured as follows.

- Section 2 outlines our methodology for developing robust econometric models to predict YWS's efficient cost allowances.
- Section 3 presents our assessment of YWS's wholesale water (WW) expenditure.
- Section 4 presents our assessment of YWS's wholesale wastewater (network plus) (WWNP) expenditure.
- Section 5 presents our assessment of YWS's bioresources (BR) expenditure.
- Section 6 presents our assessment of YWS's residential retail expenditure.

---

[24] Ofwat (2023), 'Econometric base cost models for PR24', April.

[25] YWS shared a dataset containing relevant APR data that was compiled by the industry.

[26] See Ofwat (2023), 'Econometric base cost modelling responses', available at: https://www.ofwat.gov.uk/consultation/pr24-econometric-base-cost-models-consultation/#Responses.

[27] See Oxera (2023), 'An assessment of Yorkshire Water's cost adjustment claims', September.

2 Methodology

2.1 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 modelling consultation in April 2023.[28] This includes data for all companies between 2012 and 2022 in the wholesale modelling, and data for companies between 2014 and 2022 in the residential retail modelling. We have supplemented this dataset with data from 2023, derived from an industry data share of the 2023 APRs, provided to Oxera by YWS.

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 2023 value.

- Urban rainfall excluding soil permeability (BON code BN4507): the 2023 value is based on the average value across 2018–22. Urban rainfall is volatile from year to year, so we consider that a smoothing approach is appropriate.
- Weighted average density—MSOA to LAD (BN4015) and MSOA (BN4006): the 2023 value is based on a linear extrapolation of historical data (2012–22).[29] This approach captures the overall upward trend in population density.
- Weighted average treatment size (WATS) (STWDP160): the 2023 value is set equal to its 2022 value. WATS is relatively stable over time, so we consider it appropriate to use the last year of actuals.

In order to predict YWS's efficient expenditure in AMP8, we have required forecasts of its cost drivers in that period. At this stage, where possible we have used the forecasts that YWS included in its business plan data tables. For the cost drivers that rely on external data sources (e.g. weighted average density, urban rainfall), we have extrapolated historical trends or averages.

---

[28] See Ofwat (2023), 'PR24 Cost Assessment Master Dataset, Wholesale Water Base Costs v4', April, available at https://www.ofwat.gov.uk/wp-content/uploads/2023/04/PR24-Cost-Assessment-Master-Dataset-Wholesale-Water-Base-Costs-v4.xlsx; Ofwat (2023), 'PR24 Cost Assessment Master Dataset, Wholesale Wastewater Base Costs v4', April, available at https://www.ofwat.gov.uk/wp-content/uploads/2023/04/PR24-Cost-Assessment-Master-Wholesale-Wastewater-Base-Costs-v4.xlsx; and Ofwat (2023), 'PR24 Cost Assessment Master Dataset, Residential Retail Base Costs v4', April, available at https://www.ofwat.gov.uk/wp-content/uploads/2023/04/PR24-Cost-Assessment-Master-Dataset-Residential-retail-Base-Costs-v4.xlsx.

[29] Middle layer Super Output Area (MSOA); Local Authority District (LAD).

During the modelling periods (2012–23 in wholesale, 2014–23 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. Up 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 modelling consultation, at this stage we merge the data for SWW and BWH into a single entity (SWB) in the years prior to the merger.[30]

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 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 modelling consultation analysis, we: (i) have treated HDD and SVE as new independent companies in the WW and residential retail modelling; and (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, given that the merger affects only one year 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.

2.2 Modelled expenditure

In our assessment of YKY'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. In the wholesale cost models,

---

[30] 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.

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 some enhancement activities, in line with Ofwat’s approach.

2.3 Cost driver analysis

In its PR24 modelling consultation Ofwat presented a series of models covering each of the services that water companies provide. We understand that these models are to be considered draft for consultation, and that the final models that Ofwat will use to assess companies’ efficient expenditure requirements at PR24 may differ from these models.

The industry was broadly supportive of some aspects of Ofwat’s PR24 consultation models. However, companies also argued that some of Ofwat’s modelling decisions were inappropriate and led to a deterioration in the statistical, economic and/or operational quality of the models. Therefore, we have augmented the consultation models (‘Augmented models’, as they are referred to in the remainder of this report) building upon the models presented by Ofwat at the PR24 modelling consultation that perform well against Ofwat’s modelling criteria and typically outperform its PR24 consultation models, as outlined below.

2.3.1 Wholesale water

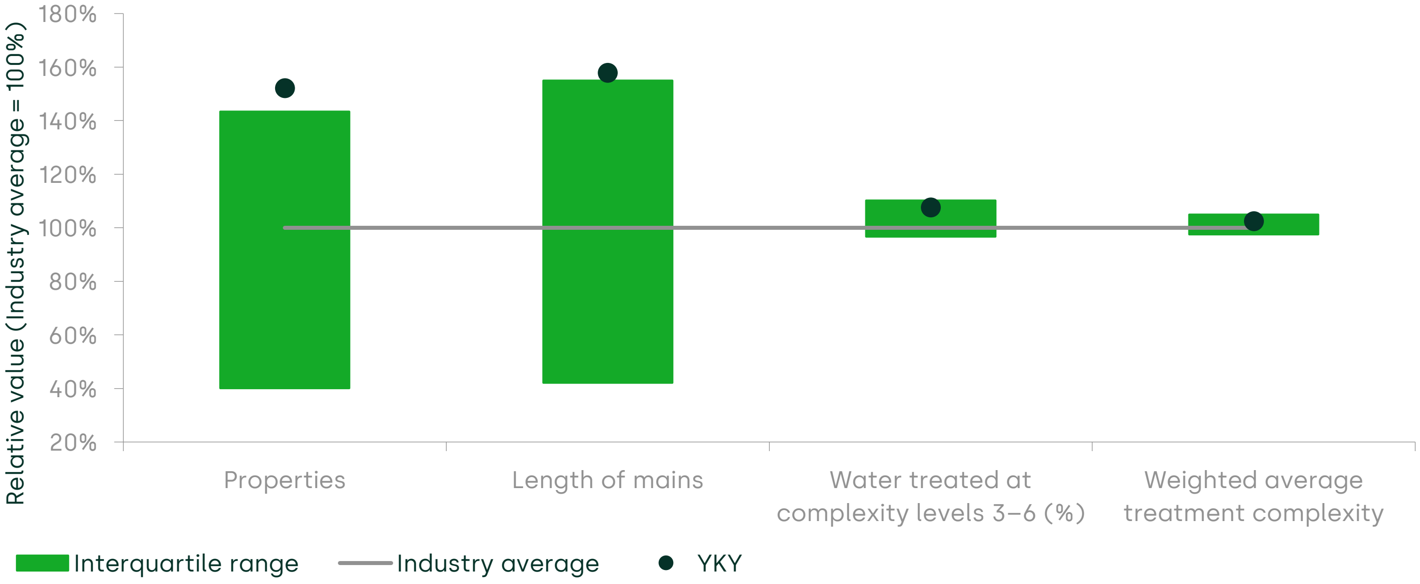
In the PR24 modelling consultation, Ofwat controlled for eight cost drivers in its proposed models. These cost drivers sought to capture the following characteristics.

- **Scale.** Ofwat controlled for connected properties as the primary measure of scale in its water resources plus (WRP) and WW models, and length of mains as the primary measure of scale in its treatment water distribution (TWD) models.
- **Density.** Ofwat presented models with three different density drivers: (i) weighted average density (MSOA to LAD); (ii) weighted average density (MSOA); and (iii) properties per length of main. In all cases, Ofwat modelled a U-shaped relationship between density and expenditure.
- **Treatment complexity.** Ofwat controlled for two measures of treatment complexity in its WRP and WW models: (i) the proportion of water treated in complexity bands W3–6; and (ii) a weighted average complexity (WAC) variable. In the case of the WAC, Ofwat modelled treatment complexity in logarithms.

- **Topography.** Ofwat controlled for two measures of pumping requirements in its TWD and WW models: (i) booster pumping stations per length of main; and (ii) average pumping head (APH) in TWD.

The figures below shows the interquartile range and how YWS compares to the industry average for the cost drivers included in Ofwat's PR24 consultation models.

Figure 2.1 Distribution of wholesale water cost drivers across the industry (2019–2023): scale and treatment complexity

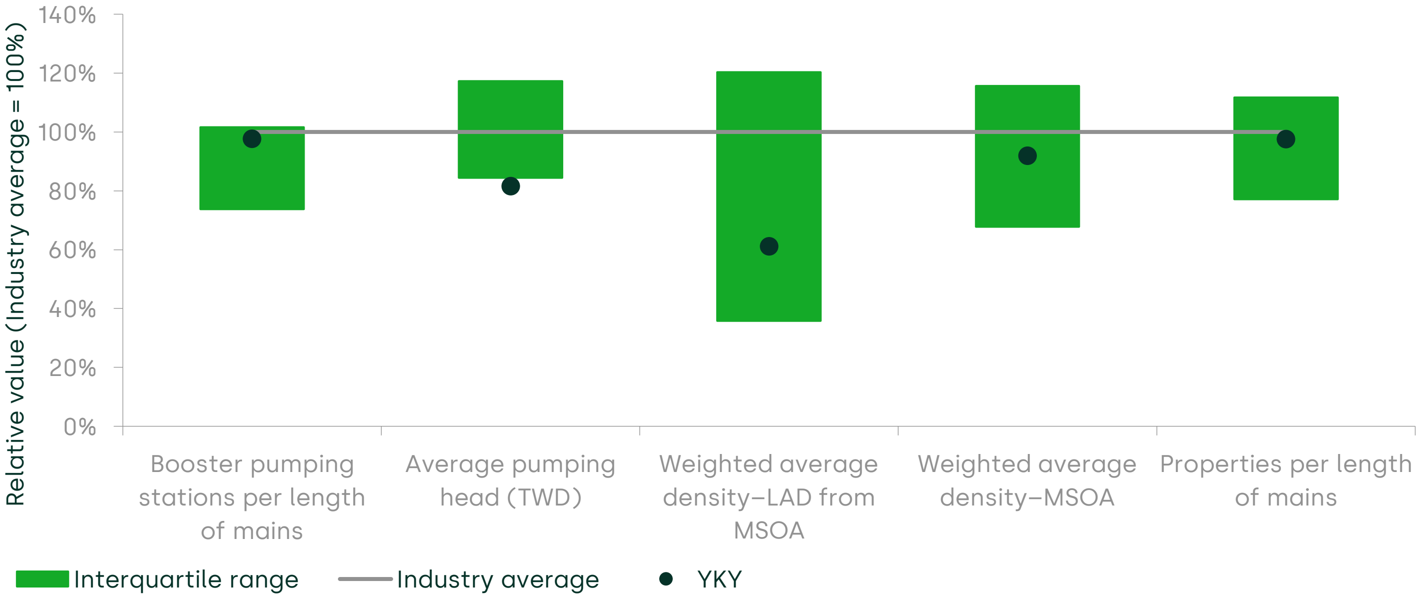


Note: To be consistent with industry-wide acronyms, we have labelled YWS as YKY.
Source: Oxera analysis.

The figure above shows that YWS is a large company in terms of both total connected properties and total length of mains (it is above the UQ). Moreover, it treats more complex water than the average company according to both water treatment complexity metrics. That is, YWS is consistently shown to be a large company that treats relatively complex water according to all of the scale and water treatment complexity variables included in Ofwat's PR24 consultation models.

The figure below shows how YWS compares to the rest of the industry on cost drivers relating to topography and population density.

Figure 2.2 Distribution of wholesale water cost drivers across the industry (2019–23): topography and density



Note: To be consistent with industry-wide acronyms, we have labelled YWS as YKY.
Source: Oxera analysis.

The figure shows that YWS's relative position with respect to topography and density varies depending on which cost driver is selected. For example, its level of booster pumping stations per length of mains is around the industry average, while it has a lower APH (TWD) than the lower quartile. Similarly, YWS is assessed to operate in a relatively sparse region according to the LAD-based weighted average density measure, but is significantly closer to the average according to the other measures of population density.

We note that the models presented by Ofwat in April's consultation represent a slight evolution of the models it used at PR19. The main differences in its current models relate to the treatment of density (where Ofwat has introduced two 'new' measures) and the treatment of topography (where Ofwat is now considering APH). The industry raised several concerns with Ofwat's PR19 models—some of which remain in the proposed models for PR24—and new concerns with Ofwat's proposed PR24 consultation models.

Therefore, we have explored and produced improved augmented models to provide a robust assessment of Yorkshire Water's performance. These models build on the industry's feedback on the PR24 consultation models (where relevant) and perform well according to Ofwat's stated modelling criteria.[31] In particular, the augmented models continue to be 'sensible and transparent', while improving upon the robustness, and/or the engineering, operational or economic rationale. The differences between the augmented models and Ofwat's PR24 consultation models are as follows.

Scale

We consider that connected properties is a reasonable measure of scale across all segments of the value chain. In relation to TWD costs (where Ofwat currently uses length of mains as the sole scale variable), connected properties may better capture the costs associated with network reinforcement than length of mains, given that one of the drivers of these costs is population growth. Indeed, the correlation between network reinforcement costs and connected properties is higher than the correlation between network reinforcement costs and length of mains (c. 0.85 versus c. 0.79).[32]

Therefore, we include connected properties as a scale variable in half of the TWD models in our augmented modelling suite, with the remaining half controlling for length of mains (as per Ofwat's PR24 consultation models).

Population density

We do not consider that properties per length of mains is a reasonable driver of density for WRP costs. Properties per length of mains relates entirely to TWD assets, and may therefore be an appropriate density driver in the TWD (and WW) models. Meanwhile, we consider that weighted average density (MSOA) may be an operationally superior measure compared to the weighted average density (MSOA to LAD), given that the former is simply a more granular measure of the latter. Therefore, we do not control for weighted average density (MSOA to LAD) in our augmented models at this stage. However, we note that the optimal granularity of the density driver will depend on exactly how density is related to costs from an operational perspective. In the absence of strong statistical evidence to support one weighted density measure over another, the selection of the most appropriate weighted average density measure may require further operational validation.

---

[31] Ofwat (2023), 'Econometric base cost models for PR24', April, p.15.

[32] Given that network reinforcement expenditure is 'lumpy' and is zero in some years for some companies, we smooth expenditure and the scale variables over the last five years before calculating these correlations. These correlations are based on the natural logarithms of the respective variables, in line with how Ofwat models expenditure in its cost models.

Treatment complexity

WAC is essentially a weighted proportion variable and should therefore be modelled in levels rather than in logarithms. Doing so allows the coefficient to be readily interpreted from an operational perspective, which is more difficult when the variable is modelled in logarithms.

Specifically, the WAC variable is the weighted sum of proportion variables, making it nearly identical to a proportion variable. The primary difference between the WAC variable and a typical proportion variable is that the former ranges from 1 to 7 while the later ranges from 0 to 1, although this difference can be simply corrected by renormalising the WAC variable without affecting the assumed relationship between treatment complexity and costs. 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 neat 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.

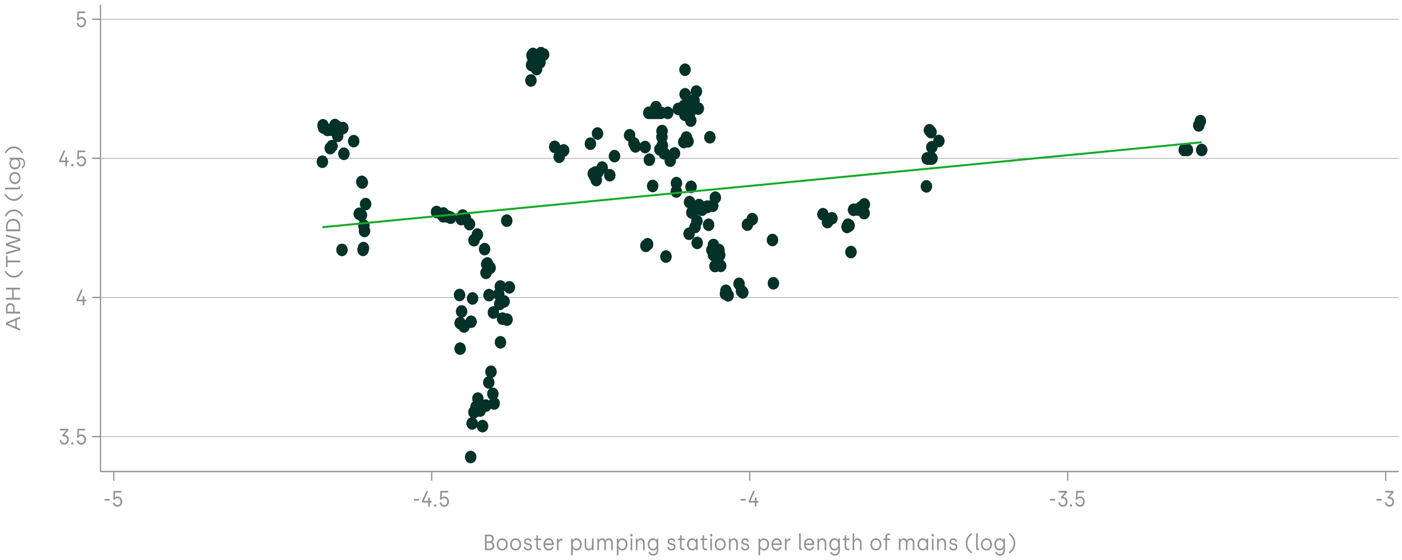
We also note that modelling this variable in levels typically improves the statistical quality of the models.

Topography

The industry raised several concerns relating to Ofwat's approach to modelling topography. Some companies argued that booster pumping stations per length of mains is not an adequate proxy for pumping costs and performs relatively poorly in the models; while other companies argued that the APH data is of a particularly poor quality.

Companies have observed that APH and booster pumping stations may capture different aspects of the costs associated with topography. For example, companies note that APH may better capture the costs of physically pumping water (OPEX), whereas booster pumping stations may better capture the maintenance and renewal costs associated with having more assets (CAPEX). Indeed, the two cost drivers are not strongly correlated with each other on the current dataset, as shown in the figure below.

Figure 2.3 Correlation between topography drivers (2012–23)



Source: Oxera analysis.

We note that the CMA attempted to include both cost drivers in the same model in the PR19 redetermination. On the data and specifications available at the time of the redetermination, the CMA found that these models performed poorly (the coefficient on APH was statistically insignificant).[33] However, on the current data and model specifications, the coefficients on the two drivers are positive and statistically significant when included in the same model.[34] That is, there is some empirical evidence based on the current dataset and model specifications that the two drivers could capture different aspects of operational costs.

Therefore, all the TWD and WW models we have considered control for both APH TWD and booster pumping stations per length of mains.[35]

2.3.2 Wholesale wastewater

In the PR24 modelling consultation, Ofwat controlled for 13 cost drivers in its proposed WWNP and BR models. These cost drivers sought to capture the following characteristics.

- **Scale.** Ofwat controlled for load as the primary measure of scale in its SWT and WWNP models. Sewer length was controlled as the measure of scale in SWC. Ofwat also controlled for sludge produced as the sole scale driver in its BR cost models.
- **Density.** Ofwat presented sewage collection cost (SWC) models with three different density drivers: (i) weighted average density (MSOA to LAD); (ii) weighted average density (MSOA); and (iii) properties per sewer length.
- **Treatment complexity.** Ofwat controlled for the proportion of load treated to ammonia consent levels ≤ 3mg/l as the treatment complexity driver across its SWT and WWNP models.
- **Topography.** Ofwat controlled for pumping capacity per sewer length as its sole topography driver across its SWC and WWNP models.
- **Economies of scale in sewage treatment.** Ofwat presented three different economies of scale in sewage treatment drivers: (i) load treated in size bands 1 to 3 (%); (ii) load treated in sewage treatment works (STWs) ≥ 100,000 people (%); and (iii) WATS.
- **Urban rainfall.** Ofwat controlled for urban rainfall per sewer length in half of its SWC and WWNP models.

---

[33] Competition and Markets Authority (2021), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report', March, para. 4.81.

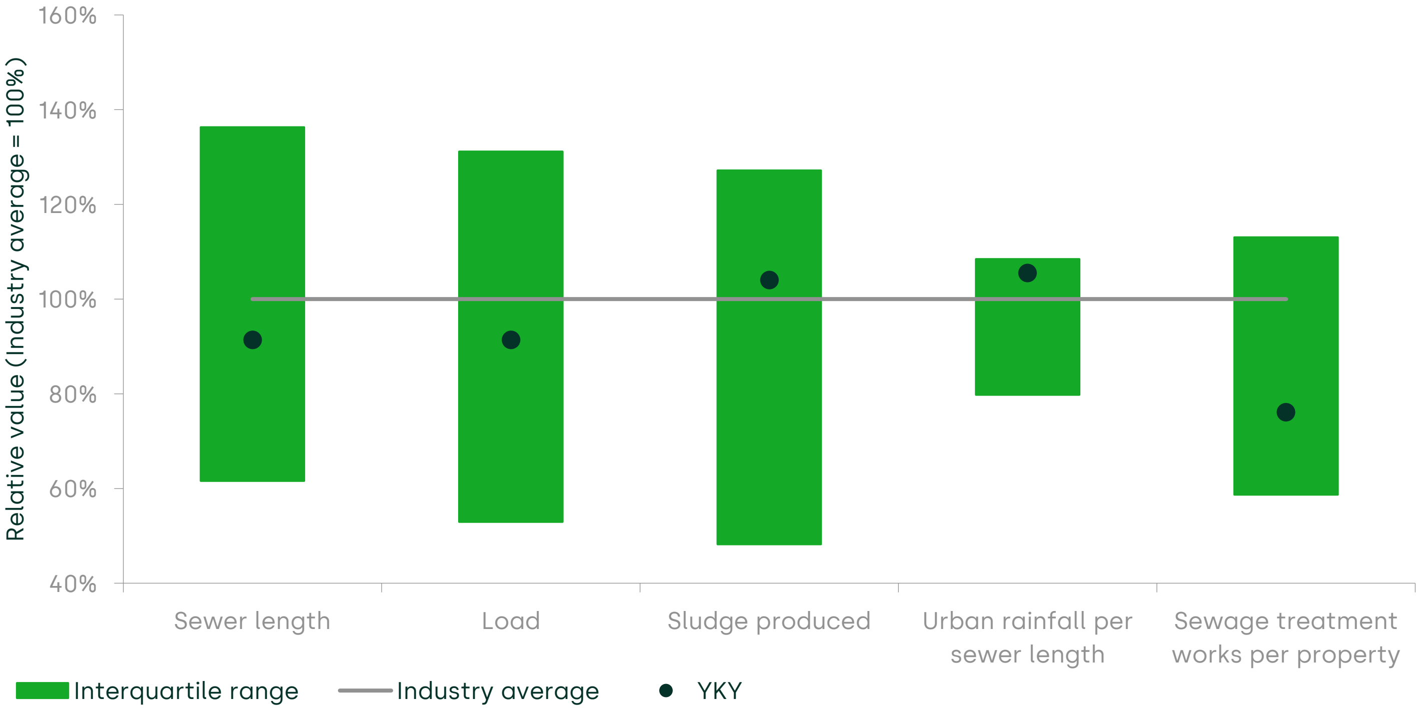
[34] Across most specifications, the significance of the coefficients on booster pumping stations and APH is largely insensitive to removing the first and last years of data, and the most and least efficient companies. However, the coefficients lose significance in some specifications under some of these sensitivities, indicating that the inclusion of both drivers in the same model may benefit from further validation operationally and as new data becomes available.

[35] We have also explored models that control for APH WRP. While the coefficient on APH WRP is statistically significant when included in the models, the coefficient on treatment complexity becomes statistically insignificant. This could indicate that APH WRP is capturing some of the costs associated with treatment complexity, rather than the costs associated with topography. Therefore, we do not include APH WRP in our proposed models.

- **Bioresources unit cost drivers.** Ofwat presented four drivers to account for economies of scale in sludge treatment: (i) load treated in bands 1–3 (%); (ii) weighted average density—LAD from MSOA; (iii) weighted average density—MSOA; and (iv) number of STWs per property.

The figure below shows the interquartile range and how YWS compares to the industry average for the cost drivers included in Ofwat's PR24 consultation models.

Figure 2.4 Distribution of wholesale wastewater cost drivers across the industry (2019–23) (i)

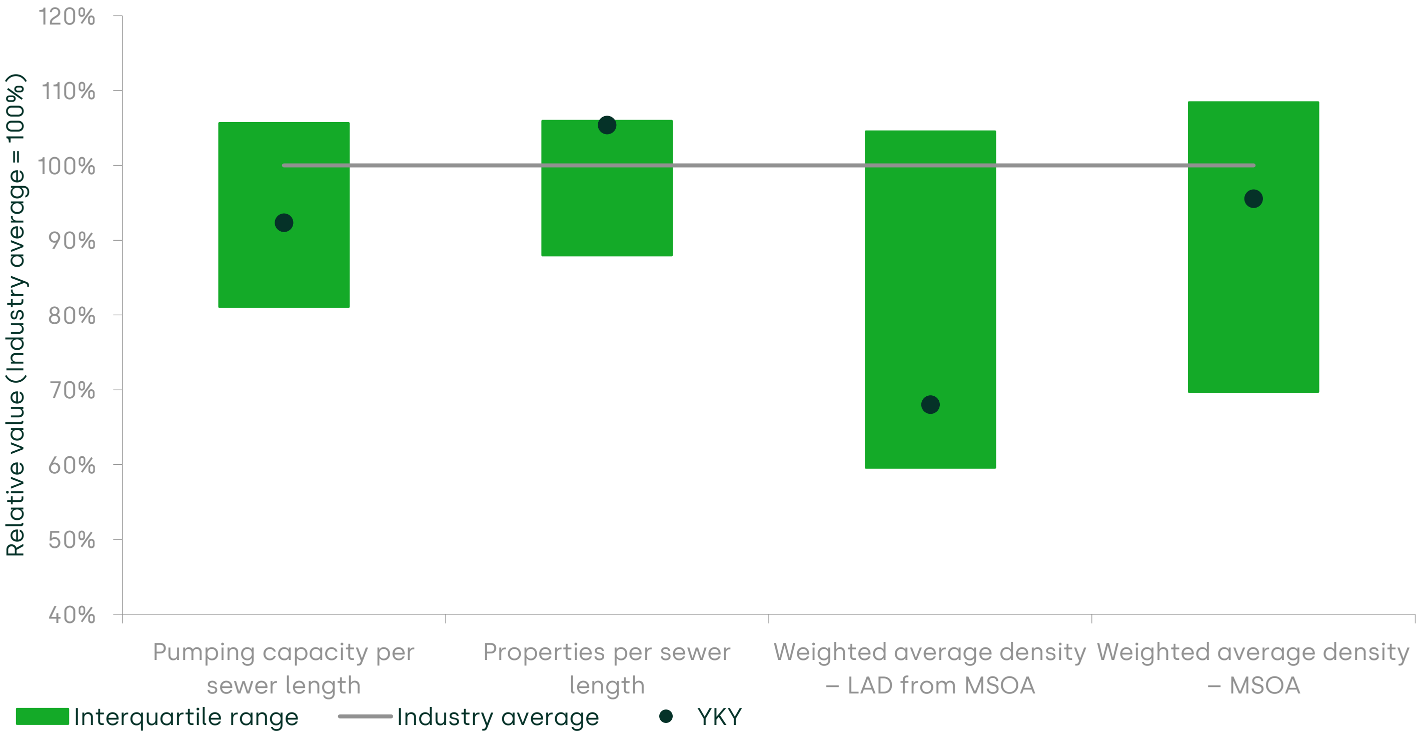


Note: To be consistent with industry-wide acronyms, we have labelled YWS as YKY.
Source: Oxera analysis.

The figure shows that the 'size' of YWS relative to that of the rest of the industry differs across different measures of scale. According to sewer length and total load, YWS is slightly smaller than the industry average; meanwhile, it is slightly larger than the industry average with respect to the total sludge produced. It operates in a region with a high degree of urban rainfall (it is close to the UQ) and also operates relatively small STWs according to the STWs per property measure.

The figure below shows how YWS compares to the rest of the industry with respect to pumping capacity and the different density measures.

Figure 2.5 Distribution of wholesale wastewater cost drivers across the industry (2019–23) (ii)

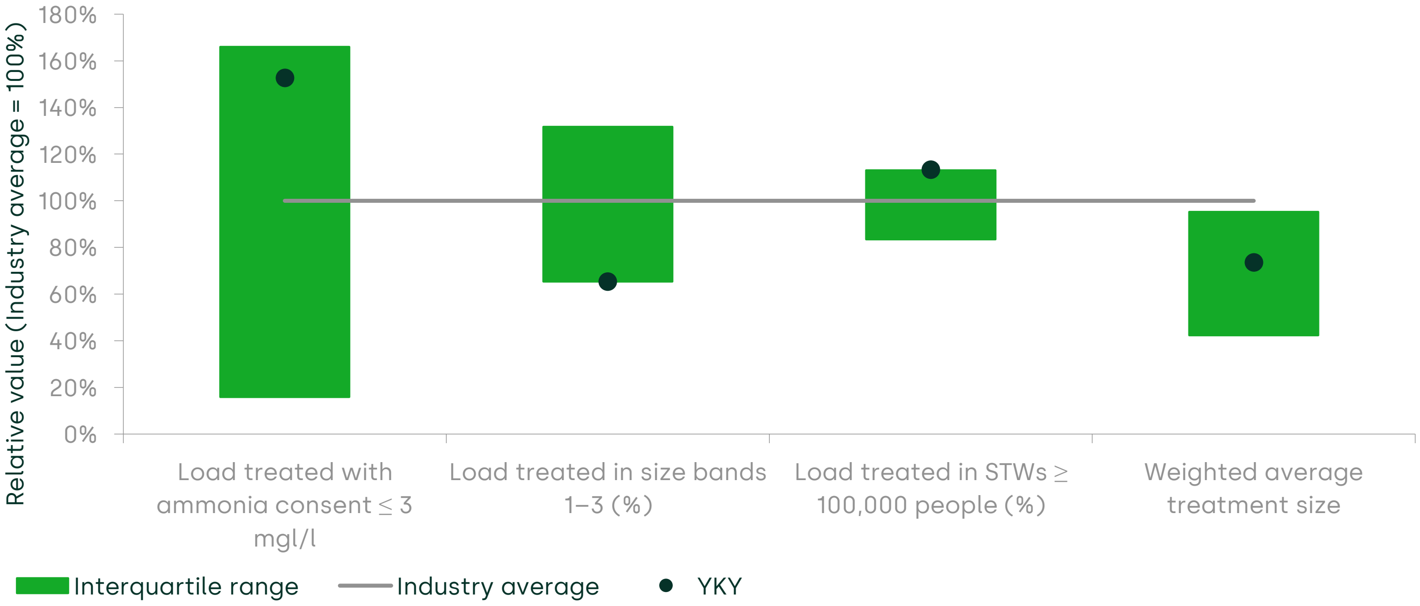


Note: To be consistent with industry-wide acronyms, we have labelled YWS as YKY.
Source: Oxera analysis.

YWS has a lower pumping capacity per sewer length than the industry average. As in WW, its relative position on population density differs across different measures of density: it is more dense than average according to the properties per sewer length measure (it is at the UQ) and is less dense than average according to the two weighted average density measures.

The figure below shows how YWS compares to the rest of the industry with respect to the treatment complexity and remaining STW-level economies of scale drivers.

Figure 2.6 Distribution of wholesale wastewater cost drivers across the industry (2019–23) (iii)



Note: To be consistent with industry-wide acronyms, we have labelled YWS as YKY.
Source: Oxera analysis.

YWS treats a relatively large proportion of its load at tight ammonia consent levels compared to the rest of the industry (it is close to the UQ). Whether or not it benefits from greater STW-level economies of scale depends on the choice of STW-level economies of scale driver assessed: YWS treats a low proportion of load at small treatment works (at the lower quartile) and a large proportion of load at large treatment works (at the UQ), but has a lower than average weighted average treatment plant size value.

We note that the models presented by Ofwat in April's consultation represent a slight evolution of the models it used at PR19. The main differences in its current consultation models relate to the treatment of density (where Ofwat has introduced two 'new' measures) and the treatment of STW-level economies of scale (where it has introduced the weighted average treatment plant size variable). The industry raised several concerns with Ofwat's PR19 models—some of which remain in the proposed models for PR24—and new concerns with Ofwat's proposed PR24 models in response to the consultation.

Therefore, we develop a suite of models—the augmented models—to provide a robust assessment of Yorkshire Water's performance. These augmented models build on the industry's feedback on the PR24 consultation models (where relevant) and perform well according to Ofwat's stated modelling criteria.[36] Note that we have not been able to develop robust BR models that are objectively superior to those presented by Ofwat in the PR24 modelling consultation on the current dataset. Therefore, we do not present augmented BR models.

The differences between our augmented models and Ofwat's PR24 consultation models are as follows.

Density

As per the augmented models in WW, we consider that the MSOA-based weighted average density measure may be operationally superior to the LAD-based measure, given that the former is simply a more granular version of the latter. Therefore, we exclude models that account for the LAD-based weighted average density variable. However, we reiterate that the optimal granularity of the density driver will depend on exactly how density is related to costs from an operational perspective.

Treatment complexity

Ofwat's treatment complexity variable accounts for the complexity associated with ammonia removal only, and does not explicitly account for alternative types of treatment complexity, such as P-removal. Given that some companies (including YWS) are anticipating a significant increase in P-removal activity in AMP8, such an omission may underestimate affected companies' efficient cost allowances. Therefore, we control for a composite treatment complexity variable in all SWT and NPWW models, which accounts for both ammonia removal and P-removal. This is discussed in more detail in our CAC report.[37]

Combined sewers

Combined sewers is an operationally relevant cost driver as such sewers require additional maintenance given that they are more prone to sewer flooding than separate sewers for foul and surface water. Their prevalence is also largely outside of management control, since they were installed before privatisation and it is difficult and costly to change the structure of a company's sewerage network. YWS has the second-highest proportion of combined sewers in the industry and is, therefore, materially affected by

---

[36] Ofwat (2023), 'Econometric base cost models for PR24', April, p.15.
[37] See Oxera (2023), 'An assessment of Yorkshire Water's cost adjustment claims', September.

Ofwat’s decision to not include combined sewers as a cost driver in its PR24 consultation models.

In the PR24 modelling consultation, Ofwat argued that the inclusion of combined sewers in the cost assessment models could ‘perversely incentivise companies not to separate sewers into surface water and foul [water]’.[38] Therefore, Ofwat proposed to use urban rainfall as a cost driver instead of combined sewers, arguing that it captures a similar impact while being more exogenous (i.e. outside the companies’ control). Ofwat’s arguments for the exclusion of combined sewers is poorly motivated, as: (i) companies cannot influence their asset base in the short run; and (ii) urban rainfall and combined sewers capture different cost pressures, and should not be considered substitutes.

On the first point, we note that Ofwat has used cost drivers that are measures of assets (and, therefore, can be influenced by companies in the long run) in its cost assessment models for successive price controls, and in the draft models that it presented in the PR24 consultation. These cost drivers include the length of the network (in both water and wastewater models), the number of booster pumping stations in the water models, and the size of treatment works in the wastewater models. In line with Ofwat’s modelling principles, it can be appropriate to control for asset-based cost drivers (including combined sewers) providing that they are exogenous in the short term.

As the cost impact of combined sewers is not currently accounted for in Ofwat’s cost assessment framework, companies are strongly incentivised to reduce the length of their combined sewers in order to perform better in Ofwat’s cost modelling (and, therefore, earn higher returns). However, the proportion of combined sewers across the industry has been static over the modelling period (2012–23).[39] That is, despite the strong incentives to reduce the length of combined sewers, companies have been unable to do so, indicating that the driver is exogenous in the short term.[40]

On the second point, Ofwat argues that the inclusion of urban rainfall (which is intended to capture costs associated with increased flooding risk) in some

---

[38] Ofwat (2023), ‘Econometric base cost models for PR24’, April, p. 45.

[39] The industry has reduced the proportion of combined sewers by only 0.02–1.60 percentage points, with an industry average reduction of 0.57 percentage points.

[40] In principle, it might be possible to test statistically whether combined sewers are exogenous in a statistical sense. Such analysis requires the identification of an ‘instrumental variable’—a variable that is correlated with combined sewers but is known to be exogenous and otherwise has no impact on companies’ costs. In principle, the proportion of combined sewers at privatisation would be a valid instrument—it is exogenous to current companies’ management and should have a strong correlation with the current level of combined sewers, given that few combined sewers have been installed since privatisation. However, such data is not publicly available.

of its PR24 consultation models means that combined sewers (which can also capture costs associated with increased flooding risk) is not required. This line of reasoning is incorrect—the observation that both drivers may capture similar characteristics does not indicate that controlling for one driver negates the need to control for the other. Indeed, this argument is inconsistent with Ofwat's P19 models, where Ofwat controlled for both population density and STW size in its bioresources models, even though both cost drivers were intended to capture the cost impact of STW-level economies of scale.[41]

Ofwat's argument rests on the assumption that an appropriate approach to model development is to group cost drivers into different categories depending on how the drivers are expected to influence costs (e.g. scale, complexity, topography), and then to select one (and only one) cost driver from each category to construct a model. In the current context, Ofwat has grouped combined sewers and urban rainfall into the same category (i.e. 'costs associated with flooding'). However, the two cost drivers could equally have been grouped into different cost categories (e.g. 'climate and weather' and 'network complexity'), in which case there would be no reason (ex ante) to exclude combined sewers from a model that controls for urban rainfall.

On the current dataset, we note that urban rainfall and combined sewers are not strongly correlated with each other.[42] Moreover, the two cost drivers perform well when included in the same model—the cost drivers are both statistically significant and are (directionally) aligned with operational expectations, and the inclusion of both drivers improves model fit. This provides empirical evidence that the two drivers capture different costs and can therefore be included in the same model.

Further justification for the inclusion of combined sewers as a cost driver can be found in our CAC report.[43]

Economies of scale

Models that control for WATS typically outperform models that control for other economies of scale drivers from a statistical perspective. The coefficient on WATS is consistently statistically significant and leads to an improved model fit.

---

[41] See Ofwat (2019), 'PR19 final determinations: Securing cost efficiency technical appendix', December, Table A2.2. Ofwat has presented similar models as part of the PR24 modelling consultation—see Ofwat (2023), 'Econometric base cost models for PR24', April, Table 7.15.

[42] The correlation coefficient is c. 0.4.

[43] See Oxera (2023), 'An assessment of Yorkshire Water's cost adjustment claims', September, section 2.

Moreover, WATS may be an operationally superior driver of economies of scale. The alternative economies of scale drivers proposed by Ofwat rely on arbitrary thresholds and assume that there is a step change in efficient costs at the STW level when these thresholds are crossed. For example, the proportion of load treated in size bands 1–3 assumes that all STWs in these size bands have the same level of efficient unit costs (e.g. STWs in size band 3 cannot benefit from additional economies of scale compared to 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 to those in size band 4). A similar logic can be applied to the proportion of load treated at STWs serving more than 100k people. We are unaware of any statistical or operational analysis that supports these assumptions.

Meanwhile, the WATS variable does not rely on arbitrary thresholds, and allows for a smoother relationship between STW size and efficient costs. Given that WATS is an operationally and statistically superior driver of STW-level economies of scale, this driver is included in all SWT and WWNP models.

Urban rainfall

As noted by Ofwat, urban rainfall is an operationally relevant driver of costs. Moreover, the cost driver is completely exogenous and performs well in the econometric models (the coefficient is consistently positive and statistically significant). While the industry has raised some concerns with the ad hoc nature of the construction of this driver, the industry was largely uncritical of the operational relevance of urban rainfall in the cost models.

Therefore, we control for urban rainfall in all of our augmented models. Going forward, it could be appropriate to explore alternative measures of urban rainfall that use more granular data. If such data became available, it may be appropriate to triangulate across models that controlled for different measures of urban rainfall, similar to how Ofwat currently triangulates across models with different measures of population density.

2.3.3 Residential retail

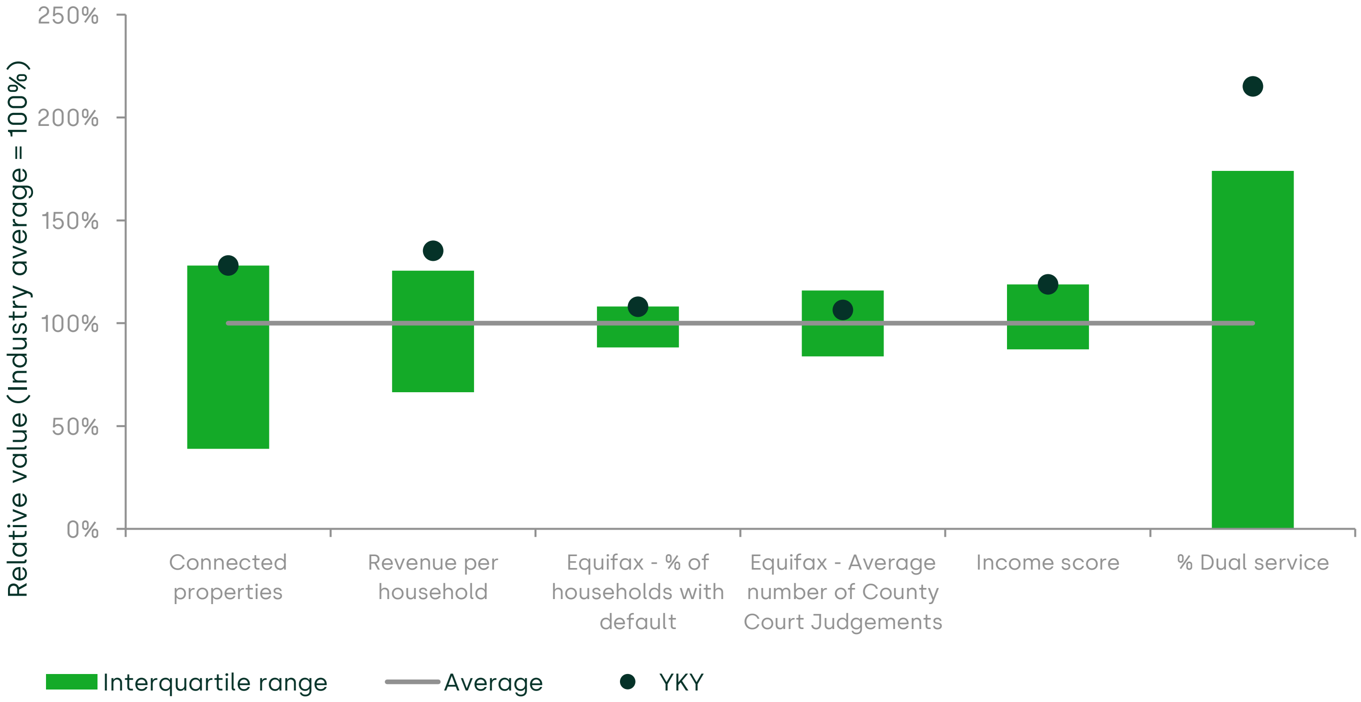
Ofwat controlled for eight cost drivers in its consultation models for residential retail. These drivers sought to capture the following characteristics.

- **Scale.** Ofwat models retail expenditure on a unit cost (expenditure per household) basis. Therefore, all models implicitly control for scale. In addition to modelling in unit cost terms, Ofwat controls for connected households in three of its RTC models and one of its ROC models to reflect economies of scale.

- **Revenue at risk.** All ROC and RTC models control for average bill size.
- **Propensity to default.** Ofwat includes three measures of deprivation to reflect a customer's propensity to default: (i) percentage of households with default; (ii) average number of County Court Judgements or Partial Insight Accounts per household; and (iii) income deprivation score.
- **Type of customer.** Ofwat controls for the proportion of dual-service households to reflect the fact that WaSCs may incur additional costs in billing water and sewerage customers (relative to WoCs that bill water customers only).
- **Time effects.** Ofwat included two time dummies, one for 2020 and one for 2021, in its RDC and RTC consultation models to account for the fact that doubtful debt (a component of bad debt costs) increased during COVID-19.

The figure below shows how YWS compares to the rest of the industry on these cost drivers.

Figure 2.7 Distribution of residential retail cost drivers across the industry (2019–23)



Note: To be consistent with industry-wide acronyms, we have labelled YWS as YKY.
Source: Oxera analysis.

The industry raised several concerns with Ofwat's proposed retail models. For example, relative to the PR19 models, Ofwat has removed cost drivers that account for population transience and the proportion of metered households. Both of these cost drivers were statistically insignificant and/or operationally unintuitive when included in the PR24 consultation models. However, given the operational relevance of these cost drivers, it may be appropriate for Ofwat to explore including these drivers in the cost models once new data becomes available.

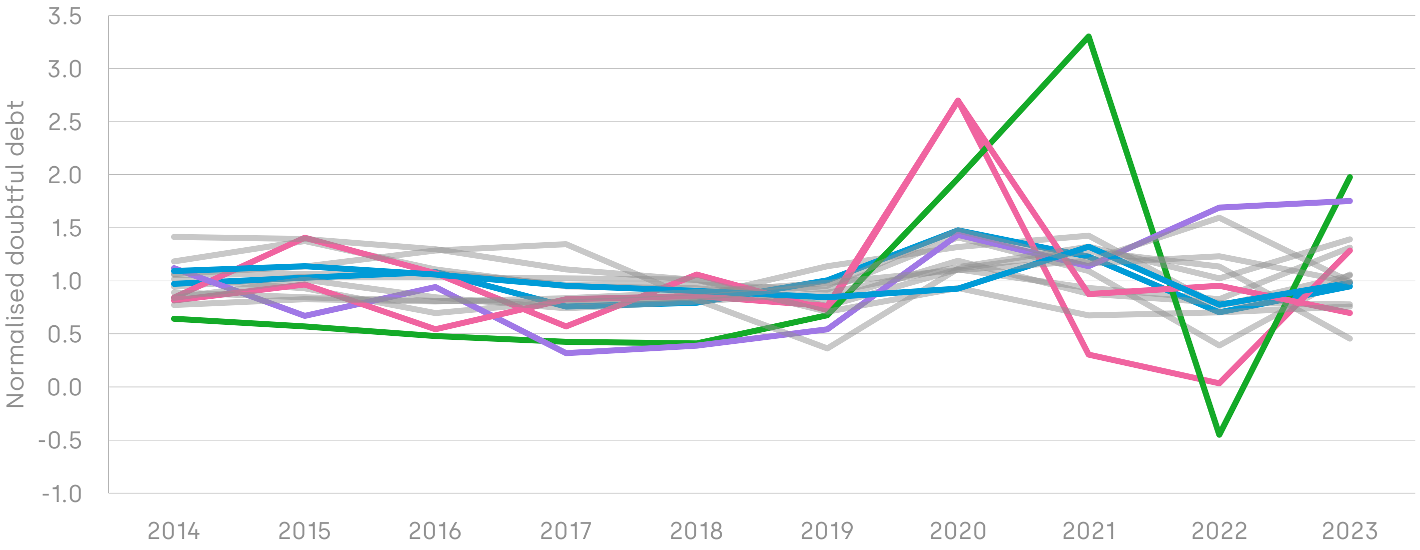
The use of time dummies to account for the increased doubtful debt costs in the COVID period is particularly concerning. The time dummies essentially remove the impact of high-cost years when determining companies' forward-looking cost allowances, assuming that Ofwat sets the dummies equal to zero in the forecast year. This assumes that there will be no event like COVID-19 in AMP8 that could cause an increase in doubtful debt, and that there are no persistent effects of COVID-19 on efficient retail costs. We are unaware of any evidence to suggest that either of these assumptions is likely.

An alternative approach could be to set the dummies equal to one in the forecast period. However, this would assume that every year of AMP8 is comparable to the COVID period, which is also unlikely. Ofwat could instead select some number between zero and one for the forecast period, which would reflect the probability of a COVID-like event occurring in AMP8 and the degree of persistence of COVID-19 effects on retail costs. However, such a decision will inevitably involve a high degree of value judgement and extensive empirical evidence.

Therefore, we have explored alternative methods for accounting for COVID-19 that mitigate the purported need for time dummies. In our augmented models, we smooth doubtful debt costs over the entire modelling period. Doing so leads to a material improvement in the statistical quality of the models in terms of model fit and statistical significance (see appendix 6.3A4).

The figure below shows how doubtful debt has evolved over the modelling period. Note that the figure shows the normalised doubtful debt costs (i.e. divided by the company's average doubtful debt costs over the modelling period) in order to make comparisons across companies.

Figure 2.8 Normalised doubtful debt (2014–23)



Source: Oxera analysis.

The figure shows that, while 2020 and 2021 might be considered unusual years for some companies, several companies (or much of the industry) did not have unusually high or low levels of doubtful debt in this period. For example, the companies' highlighted in pink (PRT and SSC) experienced a spike in doubtful debt costs in 2020, followed by a significant drop in doubtful debt in 2021. The company highlighted in green (SES) experienced a large increase in doubtful debt in 2020 and 2021, followed by a negative doubtful debt in 2022, followed by another spike in 2023. The companies highlighted in blue (WSH and WSX) have had (comparatively) stable doubtful debt costs over the entirely modelling period, and do not appear to be particularly affected by the COVID-19 years. Finally, the company highlighted in purple (SEW) has had generally higher doubtful debt costs since the COVID-19 period.

The use of COVID-19 dummies ignores the observation that COVID-19 has had widely different effects on companies' doubtful debt costs—instead, the models assume that all companies experienced higher costs in 2020 and 2021 by a fixed percentage, and that there is otherwise no unexplained volatility in doubtful debt expenditure over time. Smoothing doubtful debt costs over the modelling period can partially account for this volatility. This is consistent with Ofwat's (and other regulators') approaches to modelling expenditure that is subject to volatility, such as depreciation.

2.4 Modelling approach

As per Ofwat's PR24 modelling consultation and precedent from PR19 (and the associated CMA redetermination), we use random effects (RE) to estimate the cost models. RE cannot explicitly distinguish between statistical noise (e.g. data errors, modelling errors) and companies' inefficiency. Therefore, regulators that adopt RE (or equivalent econometric methods, such as ordinary least squares, OLS) make ad hoc adjustments to the estimated gap between companies' actual costs and the benchmark. At PR19, Ofwat set the benchmark at the fourth-ranked company in WW, the third-ranked company in WWW and the UQ company in residential retail.[44] However, at the PR19 redetermination the CMA relaxed the stringency of the benchmark to the UQ in WW and WWW, arguing that (among other things) a more stringent benchmark could not be supported by the quality of the models.

The determination of the benchmark is essentially a question of how much of the gap between companies' observed costs and predicted costs can be attributed to statistical noise versus inefficiency. Therefore, the benchmark should be determined by an assessment of the overall quality of the model which, in turn, should be driven by empirical analysis. In our assessment, we have explored the following empirical approaches to determining the benchmark.

First, we have examined the width of the confidence intervals around companies' cost predictions, which is a direct measure of the level of uncertainty in the models.[45] This technique has been considered by regulators to inform the benchmark and was also investigated by the Competition and Markets Authority (CMA) in the PR19 redetermination.[46]

While this approach has precedent and provides a measure of uncertainty based on Ofwat's own modelling assumptions, it does not in itself provide an assessment of exactly what the appropriate benchmark should be; rather, it can assess whether the uncertainty is higher or lower than the uncertainty in previous regulatory decisions, which could support whether the benchmark should be less or more stringent. That is, the method requires an anchor based on past regulatory decisions. As noted by Ofwat, the PR24 consultation models build on the PR19 models[47] and, as such, it is appropriate to compare the confidence intervals in the models outlined in this report with the confidence intervals in the PR19 models.

---

[44] In residential retail, Ofwat used a combination of a UQ benchmark based on outturn performance and a UQ benchmark based on forward-looking performance.

[45] In particular, the wider (narrower) the confidence interval, the more (less) uncertainty there is in the models. An assessment of this noise (uncertainty) to signal (inefficiency) ratio in comparison to past regulatory decisions in water and other sectors in the UK and elsewhere can be used to inform the level of benchmark or an acceptable correction for uncertainty.

[46] See CMA (2020), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report', March, pp. 202–210.

[47] Ofwat (2023), 'Econometric base cost models for PR24', April, section 1.

Second, we have employed stochastic frontier analysis (SFA) to assess the level of statistical noise (e.g. data and modelling errors) in the models. Unlike RE, SFA can explicitly decompose the residual into statistical noise and inefficiency. Moreover, SFA can be used to statistically test for the presence of inefficiency in the sample. The average inefficiency estimated by the SFA model can be compared with the average inefficiency in the RE models under different benchmarks to derive the appropriate choice of benchmark.

As this technique provides a data-driven assessment of the amount of noise in the models, it does not require the same reliance on previous regulatory decisions as the confidence interval analysis outlined above. SFA is the most commonly used econometric method for efficiency assessment by regulators across Europe[48] and has also been considered by UK regulators to assess the level of uncertainty in models.[49]

As with all econometric approaches, SFA models require assumptions regarding, for example, how the differences between predicted and actual costs (i.e. the 'residual') are distributed, and how the data is structured. In this respect, we consider two forms of SFA modelling. First, we apply a simple pooled SFA model that treats each data point as independent. This is equivalent to OLS, which was one of the estimation approaches used at PR14 and PR19, and the estimation approach used by Ofgem in the RIIO-2 controls.[50]

We also apply an SFA model that takes into account the panel structure of the dataset, equivalent to the RE estimator used by Ofwat in the modelling consultation. The specific SFA model we use in this report is often referred to as the 'four-component model'. This is because it can separate the residual into four components: (i) uncontrollable fixed differences in companies' operating environments; (ii) permanent differences in efficiency; (iii) time-varying statistical noise; and (iv) time-varying inefficiency.[51]

We note that, in the past, Ofwat has raised additional arguments for its decision relating to the benchmark. For example, it has argued that it must consider the magnitude of the efficiency challenge, and how that magnitude of the efficiency challenge compares to previous regulatory decisions. In

---

[48] For example, SFA was used by the Bundesnetzagentur (the German energy regulator) to estimate the static efficiency of German electricity distribution system operators. See Bundesnetzagentur (2018), 'Decision BK4-18-056', November.

[49] For example, ORR has used SFA to assess the efficiency of both Network Rail and Highways England. See ORR (2013), 'PR13 Efficiency Benchmarking of Network Rail using LICB', August; and ORR (2017), 'Benchmarking regional maintenance costs on England's Strategic Road Network').

[50] For example, see Ofgem (2022)', 'RIIO-ED2 Final Determinations Core Methodology Document', November.

[51] See Kumbhakar, S.C., Lien, G. and Hardaker, J.B. (2012), 'Technical efficiency in competing panel data models: A study of Norwegian grain farming', *Journal of Productivity Analysis*, **41**:2, September, pp. 1–7.

particular, Ofwat has made an argument for strengthening the efficiency challenge in successive price controls.[52]

We note that Ofwat's argument is incorrect for (at least) two reasons. First, selecting a benchmark based on the magnitude of the efficiency challenge amounts to 'goal-seeking' and defeats the purpose of undertaking a rigorous cost assessment exercise—if Ofwat already has an idea of what the appropriate challenge should be, it begs the question of why Ofwat uses econometric methods at all in its assessment. Indeed, the CMA was critical of Ofwat's argument in this regard in the PR19 redetermination, where it disagreed with Ofwat's approach to benchmark selection.[53]

Second, companies have been under incentive regulation for over three decades, during which time they have been subject to the pseudo-competitive pressure of cost benchmarking at each price review. Therefore, the scope of future efficiency gains is expected (in theory) to reduce over time. While the economic argument that the scope for efficiency improvements should reduce over time is relatively clear, this argument is not possible to test empirically under Ofwat's framework.

We note that Ofwat does not use empirical methods (such as SFA) that could support or refute this hypothesis.

2.5 Post-modelling adjustments

Neither Ofwat's PR24 consultation models nor the augmented models presented in this report can capture all of the cost pressures that companies face. As such, we also explore in this report a selection of post-modelling adjustments.

2.5.1 Cost adjustment claims

Ofwat uses econometric cost modelling to assess efficient expenditure for water companies, and has developed wholesale cost models as part of its PR24 base cost modelling consultation. These models seek to account for differences across the industry in terms of size, treatment complexity, pumping requirements, and population density.

However, it is widely acknowledged that certain factors influencing a company's expenditure may not be adequately captured by these cost models. This may result in a company appearing inefficient (or efficient) on

---

[52] See Ofwat (2020), 'Reference of the PR19 final determinations: Cost efficiency – response to common issues in companies' statements of case', May, section 6.

[53] See CMA (2021), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report', March, para. 4.493..

the basis that it suffers (or benefits) from a characteristic that is not properly accounted for.

As such, Ofwat has implemented the CAC process whereby it can make post-modelling adjustments to companies' estimated efficient expenditure requirements to reflect well-evidenced characteristics that are omitted or inappropriately reflected in the models.

YWS has commissioned us to explore CACs relating to two characteristics: (i) the prevalence of combined sewers; and (ii) the increase in P-removal activity. As noted, the assessment of YWS's CACs is provided in a separate report alongside this document. Where stated, we have included the CAC values to the modelled efficient allowances.

2.5.2 Input price uplift

Water companies and other regulated utilities are typically considered 'price takers' for the inputs that they require in the production process. That is, water companies cannot exert market power to influence the price of inputs (e.g. energy, chemicals) that are instead determined by wider market forces. Companies are incentivised to manage the impact of input prices through better planning and negotiation. As such, input prices are often considered as exogenous in the cost assessment process. For these reasons, regulators typically assess the impact of input price pressure (IPP) or real price effects (RPEs) on companies' efficient costs when determining allowed revenues.

At PR19, Ofwat commissioned Europe Economics to assess the impact of RPEs and nominal input price pressure in wholesale and retail, respectively. Once the relevant RPE was calculated, Ofwat applied the RPE from the first forecast year (at PR19, this was 2019/20). The implicit assumption under this approach is that the base cost models funded companies on the basis of the input prices faced in 2018/19, and the RPE adjustment applied to forecast data reflected the expected change in real input prices that companies would face from that year onwards.

As part of this submission, we do not comment on the appropriate estimation of an RPE or the mechanism to manage RPEs fairly over AMP8. Rather, we consider that Ofwat's assumption should be revisited—i.e. that the models fund companies on the basis of the input prices faced in the last year of the benchmarking period (expected to be either 2023/24 or 2024/25 at PR24)—given the significant and volatile input prices that companies have faced in recent years.

Specifically, if Ofwat continues to use the last five years of outturn data to estimate the cost benchmark, the cost models will fund companies on the basis of the average real input prices faced by companies in that period. In this way, the recent high input prices that the industry has faced could be 'averaged out' with outdated (and unrepresentative) data, such that the implicitly funded input prices are lower than the prices that companies face in the last year of outturn data.

We understand that Ofwat is currently collecting data from the industry in an effort to model the impact of this phenomenon, specifically relating to energy prices. However, we note that this phenomenon affects more than just energy prices; it also affects chemicals, materials and labour prices.

Given that Ofwat's data collection (and potential modelling) is currently under way, in this report we do not calculate an uplift for input prices. We look forward to working with Ofwat and the industry on this issue in future.

2.5.3 Ongoing efficiency and real price effects

The models we use to assess YWS's efficient cost requirements are based on YWS operating as an efficient company in the last five years of outturn data. However, an efficient company's costs may evolve over the upcoming regulatory period as a result of (among other factors):

- ongoing efficiency—this relates to the productivity improvements that an efficient company could make over time, driven by improved management practices and technological advancements. These productivity improvements can relate to reduced efficient expenditure for a given level (or quality) of output, increased level (or improved quality) of output for a given efficient cost, or some combination of the two;
- RPEs—as noted above, companies are considered price takers with respect to the inputs used in the production process. The efficient level of expenditure in AMP8 will depend on how these input prices evolve.

A robust assessment of the potential for ongoing efficiency improvements and RPEs is outside the scope of this report. Therefore, none of the efficient cost predictions presented in this report accounts for ongoing efficiency or RPEs.

3 Wholesale water

In this section, we present our top-down assessment of YWS's efficient base costs for wholesale water. As outlined in section 2.3.1, we use two suites of models to assess YWS's efficient costs: (i) the augmented models; and (ii) Ofwat's PR24 consultation models. These are covered in the respective sub-sections below.

3.1 Augmented models

As discussed in section 2.3.1, we have developed augmented WW models that perform well against Ofwat's criteria. The specific models are as follows.

- **WRP.** Two models that control for connected properties as a scale variable and MSOA-based weighted average density. The models differ with respect to the choice of treatment complexity variable: one model controls for the proportion of water treated in complexity bands W3–6, while the other controls for WAC.
- **TWD.** Four models that control for both APH (TWD) and booster pumping stations. The models differ with respect to the choice of scale variable (connected properties versus length of mains) and the choice of density variable (MSOA-based weighted average density and properties per length of mains).
- **WW.** Four models that control for connected properties as a scale variable and both APH (TWD) and booster pumping stations. The models differ with respect to the choice of density variable (MSOA-based weighted average density and properties per length of mains) and the choice of treatment complexity variable (proportion of water treated in complexity bands W3–6 versus WAC).

The model fit in these augmented models is typically higher than that in Ofwat's consultation models, the estimated coefficients are (directionally) aligned with operational expectations, and the coefficients are statistically significant (or close to the standard significance thresholds). These models are presented in appendix 6.3A1.1.

The table below shows the estimated confidence intervals around companies' cost predictions in these augmented models.

Table 3.1 Estimated confidence intervals in augmented wholesale water models

	WRP	TWD	WW
Augmented model 1	21%	9%	12%
Augmented model 2	23%	9%	12%
Augmented model 3		12%	12%
Augmented model 4		9%	12%
Average	**22%**	**10%**	**12%**

Note: The figures presented in the table represent the width of the 95% confidence interval around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% confidence interval ranges from -X% of the predicted costs to + X% of the predicted costs.
Source: Oxera analysis.

The table shows that the confidence intervals in these models are broadly equivalent to those of Ofwat's PR24 consultation models, despite the improvements to model quality. Therefore, even in these improved models, companies' costs are still estimated with a relatively high degree of uncertainty.

The table below compares average estimated efficiency gaps in each model under a UQ benchmark relative to SFA.

Table 3.2 Average estimated efficiency gaps (2019–23)

	Average gap to UQ			Average inefficiency SFA (pooled)		
	WRP	TWD	WW	WRP	TWD	WW
Augmented model 1	18%	9%	8%	20%*	0%	8%
Augmented model 2	20%	8%	8%	6%	6%	10%
Augmented model 3		10%	8%		3%	11%**
Augmented model 4		8%	8%		6%	13%***
Average	**19%**	**8%**	**8%**	**13%**	**4%**	**10%**

Note: The last three columns include the likelihood ratio (LR 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.

In the WRP and TWD models, the UQ benchmark always results in an efficiency challenge greater than (or equal to) the efficiency challenge implied by the pooled SFA modelling. Indeed, in half of the WRP and WW models and all of the TWD models the SFA modelling does not detect any statistically significant inefficiency. This indicates that much of the estimated efficiency gap in these augmented models is driven by statistical noise rather than inefficiency.

The panel SFA models (which account for unobserved heterogeneity across companies) suggest that there is no statistically significant inefficiency in any model, which is further evidence that much of the estimated efficiency gap in the augmented models is driven by statistical noise.

Therefore, we do not consider that the evidence supports a benchmark more stringent than the UQ, and less stringent benchmarks may be more appropriate. The table below shows YWS's efficient cost prediction under the three benchmarks considered in section 3.2.

Table 3.3 AMP8 predictions with different benchmarks (Augmented models)

	Augmented models
AMP8 predictions (average efficiency)	£1,763m
AMP8 predictions (upper-quartile efficiency)	£1,762m
AMP8 predictions (upper-tercile efficiency)	£1,775m

Note: These figures are estimated using YWS's own forecasts. The efficiency of the upper tercile is slightly above 100%, at 101%, which inflates the provision by c. £12m compared to when no frontier shift is applied.
Source: Oxera analysis.

3.2 PR24 consultation models

To the extent that the PR24 consultation models performed well using Ofwat's base modelling dataset, the models continue to perform well when outturn data for 2023 is included in the sample. The coefficients remains statistically significant and are (directionally) aligned with operational expectations, while the other model diagnostics are largely unchanged. The details of these models are given in appendix 6.3A1.2.

The table below shows the average confidence interval around companies' cost predictions in these models.

Table 3.4 Estimated confidence intervals in Ofwat's PR24 consultation models for wholesale water

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

Note: The figures presented in the table represent the width of the 95% confidence interval around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% confidence interval ranges from -X% of the predicted costs to + X% of the predicted costs.
Source: Oxera analysis. [1] Competition and Markets Authority (2021), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report', March, p. 211.

The table shows that the confidence intervals in the WRP and WW models are wider than they were at the PR19 redetermination, while those in the TWD models are narrower than they were at the PR19 redetermination (albeit still wider than the WW models at the PR19 redetermination). This indicates that the models predict companies' costs with a high degree of uncertainty—higher than the uncertainty in the PR19 redetermination.

The table 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.

Table 3.5 Average estimated efficiency gaps (2019–23)

	Average gap to UQ			Average inefficiency SFA (pooled)		
	WRP	TWD	WW	WRP	TWD	WW
Model 1	15%	10%	8%	18%*	1%	5%
Model 2	16%	9%	7%	6%	0%	6%
Model 3	16%	9%	6%	20%*	3%	7%
Model 4	17%	12%	5%	0%	8%	8%
Model 5	14%	8%	7%	21%***	0%	11%**
Model 6	16%	10%	7%	14%	9%	11%**
Model 7			12%			12%**
Model 8			14%			11%*
Model 9			12%			16%**
Model 10			13%			15%**
Model 11			9%			13%***
Model 12			9%			13%***
Average	**16%**	**10%**	**9%**	**13%**	**3%**	**11%**

Note: The last three columns include the LR 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.

The analysis shows that the average gap to the UQ is higher than the average efficiency estimated in the SFA models for both WRP and TWD. Indeed, half of the WRP models and all of the TWD models do not detect any statistically significant inefficiency in the sample, indicating that much or all of the estimated efficiency gap in Ofwat's consultation models is driven by statistical noise. In the WW models, the average efficiency gap in SFA is c. 2 percentage points higher than a UQ benchmark would suggest; however, a third of the WW models detect no statistically significant inefficiency in the sample.

While the pooled SFA models appear to show 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, we note that panel SFA models do not detect any statistically significant inefficiency in the sample in any model specification. That is, once unobserved company heterogeneity is accounted for, most (or all) of the remaining estimated efficiency gap is driven by statistical noise.

Given the analysis presented above, we do not consider that the PR24 consultation models predict companies' costs with a lower degree of uncertainty than the PR19 models, where the CMA applied a UQ benchmark. While the confidence intervals are slightly narrower in the TWD models than at PR19, the results from pooled SFA indicate that there is a significant amount of noise in the data and models. Only the WW models have some evidence that a UQ benchmark might be appropriate—the confidence intervals are slightly wider in the PR24 consultation models than at the PR19 redetermination, and the pooled SFA models show that the average efficiency gap is similar to a UQ benchmark. However, even in these models, panel SFA modelling suggests that much of the estimated efficiency gap is still driven by statistical noise.

Therefore, we consider that there is no compelling evidence to apply a benchmark more stringent than the UQ—indeed, much of the analysis suggests that a UQ benchmark may also not be supported by the evidence, and that a less stringent benchmark would be more appropriate.

The table below shows YWS's estimated cost allowance under three different benchmarks: (i) the UQ benchmark; (ii) the upper tercile (i.e. upper third) benchmark; and (iii) the average benchmark.

Table 3.6 AMP8 predictions with different benchmarks (PR24 consultation models)

	Ofwat's PR24 consultation models
AMP8 predictions (average efficiency)	£1,744m
AMP8 predictions (upper-quartile efficiency)	£1,730m
AMP8 predictions (upper-tercile efficiency)	£1,764m

Note: These figures are estimated using YWS's own forecasts. The efficiency of the upper tercile is slightly above 100%, at 101%, which inflates the provision by c. £20m compared to when no frontier shift is applied.
Source: Oxera analysis.

3.3 Summary

The table below summarises the estimated efficient AMP8 allowance under the different modelling suites considered in this section.

Table 3.7 AMP8 predictions with different benchmarks

	Augmented models	Ofwat's PR24 consultation models
AMP8 predictions (average efficiency)	£1,763m	£1,744m
AMP8 predictions (upper-quartile efficiency)	£1,762m	£1,730m
AMP8 predictions (upper-tercile efficiency)	£1,775m	£1,764m

Note: These figures are estimated using YWS's own forecasts. The efficiency of the upper tercile is slightly above 100% for both model suites, at c. 101%, which inflates the provision by c. £20m and £12m, respectively, when compared to no frontier shift applied.
Source: Oxera analysis.

The analysis suggests that YWS's efficient cost requirements in AMP8 may be between £1,730m (Ofwat's consultation models at a UQ benchmark) and £,1775m (augmented models at an upper-tercile benchmark) on the basis of these models.

We understand that YWS will be submitting CACs relating to meter renewals and increased asset maintenance and replacement activity in AMP8. An assessment of these CACs is beyond the scope of this report.

4 Wastewater network plus

In this section, we present our top-down assessment of YWS's efficient base costs for WWNP. As outlined in section 2.3, we use two suites of models to assess YWS's efficient costs: (i) The augmented models; and (ii) Ofwat's PR24 consultation models. These are covered in the respective sub-sections below.

4.1 Augmented models

As discussed in section 2.3.2, we have developed augmented wastewater models that perform well against Ofwat's criteria. The specific models are as follows.

- **SWC.** Two SWC models that control for sewer length, pumping capacity per sewer length, urban rainfall and combined sewers. The models differ with respect to the choice of density variable (MSOA-based weighted average density versus properties per sewer length).
- **SWT.** One SWT model that controls for: (i) total load; (ii) WATS; and (iii) a composite complexity variable.
- **WWNP.** One WWNP models that controls for: (i) total load; (ii) pumping capacity; (iii) WATS; (iv) combined sewers; and (v) a composite complexity variable.

The model fit in these augmented models is typically higher than Ofwat's consultation models, the estimated coefficients are (directionally) aligned with operational expectations and the coefficients are statistically significant (or close to the standard significance thresholds). These models are presented in appendix 6.3A2.1.

The table below shows the estimated confidence intervals around companies' predicted costs in the proposed wastewater models.

Table 4.1 Estimated confidence intervals in augmented wholesale wastewater models

	SWC	SWT	WWNP
Alternative model 1	12%	10%	6%
Alternative model 2	13%	-	-
Average	**12%**	**10%**	**6%**

Note: The figures presented in the table represent the width of the 95% confidence interval around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% confidence interval ranges from -X% of the predicted costs to + X% of the predicted costs.
Source: Oxera analysis.

The confidence intervals in the augmented models are narrower than in the equivalent PR24 consultation models. This suggests that the proposed improvements to the PR24 consultation models lead to more precise predictions of companies' costs, at least on an outturn basis.

The table below compares the average estimated efficiency gap under a UQ benchmark to the average efficiency gap estimated under SFA.

Table 4.2 Average estimated efficiency gaps (2019–23)

	Average gap to UQ			Average inefficiency SFA (pooled)		
	SWC	SWT	WWNP	SWC	SWT	WWNP
Alternative model 1	2%	5%	3%	5%*	6%***	5%**
Alternative model 2	4%			7%**		
Average	**3%**	**5%**	**3%**	**6%**	**6%**	**5%**

Note: The last three columns include the LR 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.

All of the augmented models detect statistically significant inefficiency, at least at the 10% level. The average efficiency gap is larger in the pooled SFA models than the UQ benchmark would suggest. However, panel SFA models do not detect any statistically significant inefficiency in the sample, suggesting that the estimated efficiency gaps under both the UQ benchmark and the pooled SFA models are driven by statistical noise.

For these reasons, we consider that a UQ benchmark may be appropriate in these augmented models, although the panel SFA models suggest that a significantly less stringent benchmark would be appropriate. The table below shows YWS's predicted efficient expenditure under the three benchmarks considered in this report.

Table 4.3 AMP8 predictions with different benchmarks (augmented models)

	Augmented models
AMP8 predictions (average efficiency)	£2,171m
AMP8 predictions (upper-quartile efficiency)	£2,131m
AMP8 predictions (upper-tercile efficiency)	£2,147m

Note: These figures are estimated using YWS's own forecasts.
Source: Oxera analysis.

4.2 PR24 consultation models

To the extent that the PR24 consultation models performed well using Ofwat's base modelling dataset, the models continue to perform well when outturn data for 2023 is included in the sample. The coefficients remains statistically significant and are (directionally) aligned with operational expectations, while the other model diagnostics are largely unchanged. The details of these models are given in appendix 6.3A2.2.

The table below shows the average confidence intervals around companies' cost predictions in these models.

Table 4.4 Estimated confidence intervals in Ofwat's PR24 consultation models for wholesale wastewater

	SWC	SWT	WWNP
Model 1	10%	17%	9%
Model 2	13%	14%	9%
Model 3	13%	10%	9%
Model 4	11%	-	7%
Model 5	11%	-	8%
Model 6	11%	-	8%
Model 7	-	-	8%
Model 8	-	-	6%
Average (PR24 models)	**12%**	**14%**	**8%**
Average (PR19)[1]	**14%**	**14.5%**	**NA**

Note: The figures presented in the table represent the width of the 95% confidence interval around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% confidence interval ranges from -X% of the predicted costs to + X% of the predicted costs.
Source: Oxera analysis. [1] CMA (2021), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report', March, p. 211.

The table shows that the confidence intervals in the PR24 consultation models are broadly comparable to the confidence intervals at the PR19 redetermination, where the CMA applied a UQ 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 PR19 models. Note that, with the exception of the WWNP models, the confidence intervals are typically wider in this part of the value chain than in the WW models.

The table below shows how the average efficiency gap in the SFA models compares to the average efficiency gap at different benchmarks, as well as whether there is any statistically significant inefficiency in the sample.

Table 4.5 Average estimated efficiency gaps (2019–23)

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

Note: The last three columns include the LR 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.

Half of the SWC models estimate that there is no statistically significant inefficiency in the sample, with the remainder of the network plus models detecting some statistically significant inefficiency, at least at the 10% significance level. The estimated efficiency gap implied by the SFA modelling in SWC and WWNP is broadly comparable to that implied by a UQ benchmark, while the average estimated efficiency gap is larger in the SWT models under SFA than under a UQ benchmark.

The confidence interval analysis and the pooled SFA models indicate that a UQ benchmark might be broadly appropriate—the confidence intervals are comparable to the PR19 redetermination (where a UQ was applied) and the pooled SFA modelling suggests that a UQ benchmark leads to a similar average efficiency challenge to SFA models. 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 once unobserved company heterogeneity is accounted for.

On the basis of the evidence above, we consider that a UQ benchmark may be broadly appropriate in the WWNP models. However, the panel SFA models indicate that the UQ benchmark may insufficiently account for the level of noise in the data and models.

The table below shows YWS's estimated cost allowance under three different benchmarks: (i) the UQ benchmark; (ii) the upper tercile benchmark; and (iii) the average benchmark.

Table 4.6 AMP8 predictions with different benchmarks (PR24 consultation models)

	Ofwat's PR24 consultation models
AMP8 predictions (average efficiency)	£1,774m
AMP8 predictions (upper-quartile efficiency)	£1,765m
AMP8 predictions (upper-tercile efficiency)	£1,794m

Note: These figures are estimated using YWS's own forecasts. The efficiency of the upper tercile is slightly above 100%, at 101%, which inflates the provision by c. £18m compared to when no frontier shift is applied.
Source: Oxera analysis.

4.3 Post-modelling adjustments

YWS is submitting two CACs relating to WWNP: one relating to the increased P-removal activity in AMP8 and the other relating to combined sewers. The details of these CACs are given in our separate CAC report.

The table below summarises YWS's efficient cost predictions for AMP8 once the CACs are accounted for. Note that the CACs are not applicable in the augmented models, given that these models already account for P-removal (through the composite complexity variable) and combined sewers.

Table 4.7 AMP8 predictions with different benchmarks including CACs

	Augmented models	**Ofwat's PR24 consultation models**
AMP8 predictions (average efficiency)	£2,171m	£1,774m
AMP8 predictions (upper-quartile efficiency)	£2,131m	£1,765m
AMP8 predictions (upper-tercile efficiency)	£2,147m	£1,794m
CAC1: Combined sewers	-	£88m
CAC2: P-removal	-	£110m
Overall AMP8 predictions	£2,131m	£1,963m

Note: These figures are estimated using YWS's own forecasts. The efficiency of the upper tercile is slightly above 100% for both model suites, at c. 101%, which inflates the provision by c. £20m and £12m, respectively, compared to when no frontier shift is applied.
Source: Oxera analysis.

5 Bioresources

In this section, we present our top-down assessment of YWS's efficient base costs for BR. Unlike for other services, we do not present augmented models in this section. While the BR models proposed in the modelling consultation have several limitations, we have not been able to develop augmented models that materially outperform Ofwat's proposed models on the existing dataset. Therefore, our top-down assessment is driven entirely by Ofwat's PR24 consultation models.

5.1 PR24 consultation models

To the extent that the PR24 consultation models performed well using Ofwat's base modelling dataset, the models continue to perform well when outturn data for 2023 is included in the sample. The coefficients remain statistically significant and are (directionally) aligned with operational expectations, while the other model diagnostics are largely unchanged. The details of these models are given in appendix 6.3A3.

The table below shows the average confidence intervals around companies' cost predictions in these models.

Table 5.1 Estimated confidence intervals in Ofwat's PR24 consultation models for bioresources

	BR total	BR unit
Model 1	23%	14%
Model 2	23%	20%
Model 3	31%	20%
Model 4	19%	20%
Model 5	29%	-
Model 6	30%	-
Average (PR24)	**26%**	**19%**
Average (PR19)	**21%**	**NA**

Note: The figures presented in the table represent the width of the 95% confidence interval around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% confidence interval ranges from -X% of the predicted costs to + X% of the predicted costs.
Source: Oxera analysis.

The estimated confidence intervals are particularly wide in the BR models when compared to the other wholesale cost models. In the BR total cost model, the average confidence interval is significantly larger than in the BR models presented at the PR19 redetermination. While the confidence intervals are narrower in the BR unit cost models compared to the total cost models, the confidence intervals are still typically wider than in other services. This suggests that the BR models predict companies costs with a high degree of uncertainty.

The table below shows how the average efficiency gap in the SFA models compares to the average efficiency gap at different benchmarks, as well as whether there is any statistically significant inefficiency in the sample.

Table 5.2 Average estimated efficiency gaps (2019–23)

	Average gap to UQ		**Average inefficiency SFA (pooled)**	
	BR total	**BR unit**	**BR total**	**BR unit**
Model 1	16%	22%	0%	0%
Model 2	19%	21%	0%	0%
Model 3	17%	19%	0%	0%
Model 4	21%	21%	0%	0%
Model 5	23%		0%	
Model 6	20%		0%	
Average	**19%**	**21%**	**0%**	**0%**

Note: The last three columns include the LR 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.

The table shows that not a single BR model detects any statistically significant inefficiency, and the average efficiency gap is 0%. This suggests that much (or all) of the estimated efficiency gap in Ofwat's consultation models is driven by statistical noise (i.e. data and modelling uncertainty) rather than inefficiency. The panel SFA models also fail to detect any statistically significant inefficiency.

For these reasons, we consider that a benchmark less stringent than the UQ should be applied, such as an average benchmark. We note that even an average benchmark may be considered too stringent if all of the estimated efficiency gap is driven by statistical noise. Nonetheless, for consistency, we present in the table below YWS's allowance under the UQ, upper-tercile and average benchmarks.

Table 5.3 AMP8 predictions with different benchmarks (PR24 consultation models)

	Ofwat's PR24 consultation models
AMP8 predictions (average efficiency)	£416m
AMP8 predictions (UQ efficiency)	£376m
AMP8 predictions (upper-tercile efficiency)	£423m

Note: These figures are estimated using YWS's own forecasts. The efficiency of the upper tercile is slightly above 100%, at 102%, which inflates the provision by c. £7m compared to when no frontier shift is applied.
Source: Oxera analysis.

We understand that YWS is not submitting a CAC in relation to BR expenditure for PR24. Therefore, the table above presents the complete top-down view of YWS's efficient AMP8 expenditure.

6 Residential retail

In this section, we present our top-down assessment of YWS's efficient base costs for residential retail. As outlined in section 2.3, we use two suites of models to assess its efficient costs: (i) the augmented models; and (ii) Ofwat's PR24 consultation models. These are covered in the respective sub-sections below.

6.1 Augmented models

As discussed in section 2.3.3, we have not amended the cost driver specification in our augmented models. Rather, we have smoothed doubtful debt costs over the modelling period and removed the time dummies that were intended to capture the increased doubtful debt provisions associated with COVID-19. Doing so leads to an improvement in the statistical quality of the models (see appendix 6.3A4.1 for details).

The table below shows the average width of the confidence intervals around companies' cost predictions in the augmented retail models.

Table 6.1 Estimated confidence intervals in augmented residential retail models

	RDC	ROC	RTC
Model 1	20%	8%	12%
Model 2	20%	9%	12%
Model 3	20%		13%
Model 4			11%
Model 5			11%
Model 6			12%
Average (augmented models)	**20%**	**9%**	**12%**

Note: The figures presented in the table represent the width of the 95% confidence interval around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% confidence interval ranges from -X% of the predicted costs to + X% of the predicted costs.
Source: Oxera analysis.

While the confidence intervals in the augmented RDC models remain wider than in the ROC and RTC models, the width has materially reduced compared to Ofwat's PR24 consultation models. There is also a reduction in the width of the confidence intervals in the RTC models, albeit to a lesser extent. That is, smoothing doubtful debt expenditure allows the models to predict companies' expenditure with less uncertainty.

The table below shows the average estimated efficiency gap in these augmented models under a UQ benchmark and under pooled SFA.

Table 6.2 Average estimated efficiency gaps (2019–23)

	Average gap to UQ			Average inefficiency SFA (pooled)		
	RDC	**ROC**	**RTC**	**RDC**	**ROC**	**RTC**
Model 1	22%	8%	15%	0%	9%**	8%***
Model 2	22%	8%	13%	0%	9%**	8%***
Model 3	21%		15%	0%		10%***
Model 4			16%			10%***
Model 5			14%			10%***
Model 6			15%			11%***
Average	**22%**	**8%**	**15%**	**0%**	**9%**	**10%**

Note: The last three columns include the LR 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.

The table shows that no RDC model detects any statistically significant inefficiency, further supporting the observation that these models estimate costs with a high degree of uncertainty. Meanwhile, the ROC and RTC models detect statistically significant inefficiency. While the efficiency gaps under SFA and the UQ benchmark are comparable in the ROC models, the estimated efficiency gap is smaller under SFA in the RTC models. Moreover, 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 presented above suggests that a benchmark less stringent than the UQ may be appropriate in these augmented models. The table below shows YWS's efficient cost predictions in these augmented models under the three different benchmarks considered in this report.

Table 6.3 AMP8 predictions with different benchmarks (Augmented models)

	Augmented models
AMP8 predictions (average efficiency)	£494m
AMP8 predictions (upper-quartile efficiency)	£467m
AMP8 predictions (upper-tercile efficiency)	£469m

Note: These figures are estimated using YWS's own forecasts.
Source: Oxera analysis.

6.2 PR24 consultation models

To the extent that the PR24 consultation models performed well using Ofwat's base modelling dataset, the models continue to perform well when outturn data for 2023 is included in the sample. The coefficients remain statistically significant and are (directionally) aligned with operational expectations, while the other model diagnostics are largely unchanged. The details of these models are given in appendix A4.2.

The table below shows the average confidence intervals around companies' cost predictions in these models. Note that the CMA did not assess retail models at the PR19 redetermination.

Table 6.4 Estimated confidence intervals in Ofwat's PR24 consultation models for residential retail

	RDC	ROC	RTC
Model 1	18%	8%	10%
Model 2	19%	9%	11%
Model 3	18%		12%
Model 4			10%
Model 5			11%
Model 6			11%
Average (PR24 models)	**18%**	**9%**	**11%**

Note: The figures presented in the table represent the width of the 95% confidence interval around companies' cost predictions and should be interpreted in +/- terms—i.e. a value of 'X%' would suggest that the 95% confidence interval ranges from -X% of the predicted costs to + X% of the predicted costs.
Source: Oxera analysis.

The table shows that the average confidence interval is significantly wider in the RDC models than in the ROC or RTC models. Indeed, the confidence intervals are wider in the RDC models than all wholesale models with the exception of BR. Meanwhile, the confidence intervals in the ROC and RTC models are comparatively narrow. This suggests that bad debt costs are estimated with a high degree of uncertainty, while other costs and total costs are estimated with a lower degree of uncertainty.

The table below shows how the estimated efficiency gap under a UQ benchmark compares to the average efficiency gap under SFA.

Table 6.5 Average estimated efficiency gaps (2019–23)

	Average gap to UQ			Average inefficiency SFA (pooled)		
	RDC	**ROC**	**RTC**	**RDC**	**ROC**	**RTC**
Model 1	17%	8%	7%	0%	9%**	10%***
Model 2	19%	8%	7%	0%	9%**	10%***
Model 3	24%		8%	0%		12%***
Model 4			10%			11%***
Model 5			12%			10%***
Model 6			10%			11%***
Average	**20%**	**8%**	**9%**	**0%**	**9%**	**11%**

Note: The last three columns include the LR 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.

The table shows that no RDC model detects any statistically significant inefficiency, further supporting the observation that these models estimate costs with a high degree of uncertainty. Meanwhile, the ROC and RTC models detect statistically significant inefficiency, and the estimated efficiency gap in these SFA models is comparable to (or greater than) that implied by a 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 confidence interval analysis and the pooled SFA models may support a UQ benchmark in the ROC and RTC models, the RDC models are estimated with materially higher uncertainty and the panel SFA models suggest that the estimated efficiency gap in all models (RDC, ROC and RTC) is driven by statistical noise rather than inefficiency.

The table below shows YWS's efficient cost predictions in the PR24 consultation retail models under the three benchmarks considered in this report.

Table 6.6 AMP8 predictions with different benchmarks (PR24 consultation models)

	Ofwat's PR24 consultation models
AMP8 predictions (average efficiency)	£491m
AMP8 predictions (upper-quartile efficiency)	£447m
AMP8 predictions (upper-tercile efficiency)	£462m

Note: These figures are estimated using YWS's own forecasts.
Source: Oxera analysis.

6.3 Summary

The table below summarises the estimated efficient AMP8 allowance under the different modelling suites considered in this section.

Table 6.7 AMP8 predictions with different benchmarks

	Augmented models	Ofwat's PR24 consultation models
AMP8 predictions (average efficiency)	£494m	£491m
AMP8 predictions (upper-quartile efficiency)	£467m	£447m
AMP8 predictions (upper-tercile efficiency)	£469m	£462m

Note: These figures are estimated using YWS's own forecasts.
Source: Oxera analysis.

The analysis suggests that YWS's efficient cost requirements in AMP8 may be between £447m (Ofwat's consultation models at a UQ benchmark) and £469m (Augmented models at an upper-tercile benchmark) on the basis of these models.

A1 Wholesale water models

The following sections set out the WW model results from Oxera's assessment, and Ofwat's provisional PR24 consultation models updated with 2022/23 data.

A1.1 Augmented models

The table below presents the model results from Oxera's assessment for WRP BOTEX and TWD BOTEX+.

Table A1.1 Augmented models for WRP and TWD

	WRPAug1	WRPAug2	TWDAug1	TWDAug2	TWDAug3	TWDAug4
Connected properties (log)	1.056***	1.050***			1.058***	1.061***
	(0)	(0)			(0)	(0)
Water treated at complexity levels 3 to 6 (%)	0.00468***					
	(0.00159)					
Weighted average treatment complexity		0.134				
		(0.133)				
Length of mains (log)			1.017***	1.061***		
			(0)	(0)		
Booster pumping stations per length of mains (log)			0.287***	0.364***	0.493***	0.364***
			(0.00454)	(0.00222)	(0.000525)	(0.00222)
Average pumping head TWD (log)			0.366***	0.326***	0.362***	0.326***
			(2.92e-09)	(3.74e-06)	(1.13e-07)	(3.74e-06)
Weighted average density—MSOA (log)	-5.204**	-4.832**	-5.957***		-6.024***	
	(0.0118)	(0.0346)	(0)		(2.55e-07)	
Weighted average density—MSOA (log) squared	0.319**	0.293**	0.416***		0.393***	
	(0.0117)	(0.0357)	(0)		(1.41e-08)	
Properties per length of mains (log)				-14.77***		-15.83***
				(0)		(0)
Properties per length of mains (log) squared				1.874***		1.874***
				(0)		(0)
Constant	10.11	8.611	15.19***	23.03***	13.13***	23.03***
	(0.186)	(0.322)	(8.73e-11)	(0)	(0.00514)	(0)
Observations	204	204	204	204	204	204
Model fit	0.897	0.896	0.968	0.971	0.968	0.971
RESET	0.824	0.67	0.739	0.714	0.541	0.714

	WRPAug1	WRPAug2	TWDAug1	TWDAug2	TWDAug3	TWDAug4
BP	0	0	0	0	0	0
VIF	490.931	515.289	503.556	714.927	504.272	715.327
Pooling	1	1	0.543	0.463	0.18	0.463
Normality	0.128	0.398	0.328	0.629	0.893	0.629
Heteroscedasticity	0	0	0.64	0.632	0.215	0.632

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

The table below presents the model results from Oxera's assessment for WW BOTEX+.

Table A1.2 Augmented models for WW

	WWAug1	WWAug2	WWAug3	WWAug4
Connected properties (log)	1.046***	1.037***	1.037***	1.028***
	(0)	(0)	(0)	(0)
Water treated at complexity levels 3 to 6 (%)	0.00293**		0.00340***	
	(0.0224)		(0.00475)	
Weighted average treatment complexity		0.109**		0.119***
		(0.0131)		(0.00377)
Booster pumping stations per length of mains (log)	0.383***	0.383***	0.282*	0.276**
	(0.00588)	(0.00493)	(0.0515)	(0.0471)
Average pumping head TWD (log)	0.288***	0.259***	0.243**	0.215**
	(0.00762)	(0.00912)	(0.0370)	(0.0439)
Weighted average density—MSOA (log)	-5.216***	-4.567***		
	(0)	(3.56e-09)		
Weighted average density—MSOA (log) squared	0.335***	0.292***		
	(0)	(3.02e-10)		
Properties per length of mains (log)			-11.31***	-10.01***
			(1.38e-08)	(3.16e-07)
Properties per length of mains (log) squared			1.325***	1.166***

	WWAug1	WWAug2	WWAug3	WWAug4
			(6.29e-09)	(1.57e-07)
Constant	10.73***	8.262**	14.39***	11.72***
	(0.000615)	(0.0113)	(0.000832)	(0.00549)
Observations	204	204	204	204
Model fit	0.967	0.969	0.967	0.970
RESET	0.586	0.594	0.721	0.402
BP	0	0	0	0
VIF	518.424	554.984	717.343	729.166
Pooling	0.996	0.993	0.985	0.988
Normality	0.628	0.479	0.067	0.073
Heteroscedasticity	0	0	0	0

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

A1.2 PR24 consultation models

The table below presents the model results for Ofwat’s PR24 provisional models, including 2022/23 data, for WRP BOTEX.

Table A1.3 Regression results for Ofwat’s PR24 for WRP

	WRP1	WRP2	WRP3	WRP4	WRP5	WRP6
Connected properties (log)	1.083***	1.077***	1.056***	1.053***	1.029***	1.024***
	(0)	(0)	(0)	(0)	(0)	(0)
Proportion of water treated in complexity bands W3–6	0.00519***		0.00468***		0.00528***	
	(0.000365)		(0.00159)		(6.51e-05)	
Weighted average density (MSOA to LAD) (log)	-1.646***	-1.536**				
	(0.00319)	(0.0150)				
Weighted average density (MSOA to LAD) (log), squared	0.105***	0.0964**				
	(0.00346)	(0.0169)				
Weighted average complexity (log)		0.440		0.417		0.462*
		(0.116)		(0.138)		(0.0848)
Weighted average density (MSOA) (log)			-5.204**	-5.114**		
			(0.0118)	(0.0254)		

	WRP1	WRP2	WRP3	WRP4	WRP5	WRP6
Connected properties (log)	1.083***	1.077***	1.056***	1.053***	1.029***	1.024***
	(0)	(0)	(0)	(0)	(0)	(0)
Weighted average density (MSOA) (log), squared			0.319**	0.312**		
			(0.0117)	(0.0249)		
Properties per length of mains (log)					-7.942**	-7.470**
					(0.0114)	(0.0190)
Properties per length of mains (log), squared					0.883**	0.824**
					(0.0165)	(0.0254)
Constant	-5.105***	-5.620***	10.11	9.617	7.051	5.915
	(0.000624)	(0.00219)	(0.186)	(0.266)	(0.278)	(0.373)
Observations	204	204	204	204	204	204
Model fit	0.906	0.901	0.897	0.894	0.906	0.902
RESET	5.51e-07	1.53e-07	0.0621	0.0128	0.00281	0.000727
BP	0	0	0	0	0	0
VIF	202.995	203.328	490.931	510.256	678.728	679.631
Pooling	1	1	1	1	0.993	0.999
Normality	0.116	0.643	0.128	0.367	0.026	0.248
Heteroscedasticity	0	0	0	0	0	0

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

As shown above, the model fit remains largely the same, albeit with a marginal deterioration in WRP4 compared to their equivalent augmented models. The p-value of the weighted average treatment complexity is also further from the 10% significance threshold than under the augmented models.

The following table presents the regression results for Ofwat's PR24 provisional models, including 2022/23 data, for TWD BOTEX+.

Table A1.4 Regression results for Ofwat's PR24 for TWD

	TWD1	TWD2	TWD3	TWD4	TWD5	TWD6
Weighted average density (MSOA to LAD) (log)	-2.857***			-3.084***		
	(6.43e-09)			(0)		
	0.228***			0.237***		

	TWD1	TWD2	TWD3	TWD4	TWD5	TWD6
Weighted average density (MSOA to LAD) (log), squared	(0)			(0)		
Weighted average density (MSOA) (log)		-5.849***			-6.722***	
		(3.26e-06)			(0)	
Weighted average density (MSOA) (log), squared		0.412***			0.458***	
		(6.21e-08)			(0)	
Properties per length of mains (log)			-15.60***			-17.15***
			(0)			(0)
Properties per length of mains (log), squared			1.979***			2.120***
			(0)			(0)
Length of mains (log)	1.074***	1.026***	1.073***	1.065***	1.018***	1.048***
	(0)	(0)	(0)	(0)	(0)	(0)
Booster pumping stations per length of mains (log)	0.403***	0.400***	0.453***			
	(0.00848)	(0.00280)	(0.00294)			
Average pumping head TWD (log)				0.334***	0.392***	0.342***
				(1.08e-05)	(1.26e-10)	(1.58e-05)
Constant	4.351***	16.58***	26.35***	2.412	17.31***	27.25***
	(0.00428)	(0.000561)	(9.09e-10)	(0.137)	(2.63e-06)	(0)
Observations	204	204	204	204	204	204
Model fit	0.955	0.952	0.958	0.960	0.964	0.965
RESET	0	0	0	0.000841	0.0183	0.647
BP	0	0	0	0	0	0
VIF	206.995	496.376	714.842	205.697	487.006	674.189
Pooling	0.73	0.839	0.871	0.694	0.652	0.782
Normality	0.038	0.012	0.894	0.662	0.972	0.55
Heteroscedasticity	0.108	0.019	0.001	0.619	0.913	0.333

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

The range of model fits under the PR24 consultation models (0.952–0.965) is lower than that of the augmented models (0.968–0.971). This indicates that the PR24 consultation models cannot explain as much of the variability in TWD expenditure than the augmented models.

The following tables present the regression results for Ofwat's PR24 provisional models, including 2022/23 data, for WW BOTEX+.

Table A1.5 Regression results for Ofwat's PR24 for WW (Part 1)

	WW1	WW2	WW3	WW4	WW5	WW6
Connected properties (log)	1.075***	1.062***	1.052***	1.043***	1.044***	1.035***
	(0)	(0)	(0)	(0)	(0)	(0)
Proportion of water treated in complexity bands W3–6	0.00340***		0.00301**		0.00351***	
	(0.00208)		(0.0107)		(0.000892)	
Weighted average density (MSOA to LAD) (log)	-1.929***	-1.694***				
	(0.000110)	(0.000395)				
Weighted average density (MSOA to LAD) (log), squared	0.137***	0.120***				
	(3.48e-05)	(0.000162)				
Weighted average complexity (log)		0.405**		0.376**		0.414**
		(0.0221)		(0.0339)		(0.0103)
Weighted average density (MSOA) (log)			-4.913***	-4.410***		
			(0.000212)	(0.000816)		
Weighted average density (MSOA) (log), squared			0.316***	0.283***		
			(7.73e-05)	(0.000375)		
Properties per length of mains (log)					-11.62***	-10.57***
					(3.70e-07)	(1.47e-06)
Properties per length of mains (log), squared					1.361***	1.231***
					(4.01e-07)	(1.50e-06)
Booster pumping stations per length of mains (log)	0.404**	0.405**	0.465***	0.456***	0.339*	0.324**
	(0.0196)	(0.0109)	(0.00597)	(0.00465)	(0.0523)	(0.0484)
Constant	-1.916	-2.853*	11.00**	8.878*	16.25***	13.85***
	(0.213)	(0.0606)	(0.0309)	(0.0837)	(0.000623)	(0.00244)
Observations	204	204	204	204	204	204
Model fit	0.964	0.966	0.961	0.964	0.964	0.966
RESET	0	0	0	6.35e-09	3.45e-09	5.77e-11
BP	0	0	0	0	0	0
VIF	205.049	205.538	506.035	526.086	717.341	721.587
Pooling	0.948	0.917	0.987	0.978	0.969	0.977
Normality	0.174	0.429	0.41	0.449	0.129	0.182
Heteroscedasticity	0	0	0	0	0	0

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

Table A1.6 Regression results for Ofwat's PR24 for WW (Part 2)

	WW7	WW8	WW9	WW10	WW11	WW12
Connected properties (log)	1.070***	1.060***	1.041***	1.035***	1.026***	1.020***
	(0)	(0)	(0)	(0)	(0)	(0)
Proportion of water treated in complexity bands W3–6	0.00311**		0.00264**		0.00324**	
	(0.0204)		(0.0496)		(0.0101)	
Weighted average density (MSOA to LAD) (log)	-2.283***	-2.110***				
	(8.17e-08)	(4.62e-07)				
Weighted average density (MSOA to LAD) (log), squared	0.156***	0.143***				
	(4.29e-08)	(2.65e-07)				
Weighted average complexity (log)		0.351*		0.324*		0.373**
		(0.0688)		(0.0875)		(0.0353)
Weighted average density (MSOA) (log)			-6.379***	-6.035***		
			(2.46e-07)	(8.28e-07)		
Weighted average density (MSOA) (log), squared			0.401***	0.378***		
			(1.85e-07)	(4.66e-07)		
Properties per length of mains (log)					-13.10***	-12.27***
					(0)	(0)
Properties per length of mains (log), squared					1.510***	1.409***
					(0)	(0)
Average pumping head TWD (log)	0.328***	0.315***	0.340***	0.328***	0.270**	0.254**
	(0.00195)	(0.00271)	(0.00356)	(0.00459)	(0.0254)	(0.0367)
Constant	-3.407*	-4.070**	14.05***	12.63**	17.51***	15.69***
	(0.0549)	(0.0231)	(0.00495)	(0.0130)	(3.64e-07)	(2.29e-06)
Observations	204	204	204	204	204	204
Model fit	0.963	0.964	0.959	0.960	0.964	0.965
RESET	7.25e-06	3.85e-06	0.00671	0.00278	0.00171	0.000290
BP	0	0	0	0	0	0
VIF	203.824	205.806	498.917	527.144	679.281	679.756
Pooling	0.88	0.902	0.976	0.976	0.984	0.987
Normality	0.031	0.186	0.08	0.139	0.005	0.026
Heteroscedasticity	0	0	0	0	0.001	0.001

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

The range of model fits across the PR24 WW models (0.959–0.966) is lower than that of the augmented WW models (0.967–0.970). This indicates that the

PR24 consultation models do not explain as much of the variability in WW expenditure as the augmented models do. Furthermore, the significance of the WAC variable is lower under the PR24 consultation models than in the augmented.

A2 Wholesale wastewater (network plus) models

The following sections set out the wastewater network plus model results from Oxera's assessment, and Ofwat's provisional PR24 consultation models updated with 22/23 data.

A2.1 Augmented models

The table below presents the augmented models for SWC, SWT, and WWNP.

Table A2.1 Alternative models for wholesale wastewater

	SWCAug1	SWCAug2	SWTAug1	WWNPAug1
Sewer length (log)	0.863***	0.919***		
	(0)	(0)		
Pumping capacity per sewer length (log)	0.404***	0.604***		0.351***
	(0.00686)	(0.000193)		(1.87e-05)
Properties per sewer length (log)	1.088***			
	(7.07e-05)			
Weighted average density—MSOA (log)		0.469***		
		(2.27e-05)		
Urban rainfall per sewer length (log)	0.0918***	0.129***		0.0651**
	(0.000443)	(0.00190)		(0.0256)
Load (log)			0.785***	0.754***
			(0)	(0)
Weighted average treatment size (log)			-0.255***	-0.0959***
			(6.91e-07)	(3.71e-05)
% Combined sewers	0.00202	0.00437*		0.00223**
	(0.140)	(0.0898)		(0.0324)
Composite complexity variable			0.00660***	0.00552***
			(1.21e-08)	(0)
Constant	-8.582***	-8.889***	-2.838***	-3.311***
	(0)	(1.34e-08)	(0.000168)	(0)
Observations	120	120	120	120
Model fit	0.920	0.920	0.900	0.961
RESET	7.03e-08	0.0848	0.750	0.0603
BP test	1.02e-05	4.37e-08	0	0.113

	SWCAug1	SWCAug2	SWTAug1	WWNPAug1
VIF	3.039	2.364	4.390	5.142
Pooling	0.935	0.971	0.964	0.825
Normality	0.209	0.0435	0.00524	0.00514
Heteroscedasticity	0.818	0.223	0.388	0.222

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

A2.2 PR24 consultation models

The table below presents the model results from Ofwat's PR24 consultation models for SWC and SWT.

Table A2.2 Regression results for Ofwat's PR24 for SWC and SWT

	SWC1	SWC2	SWC3	SWC4	SWC5	SWC6	SWT1	SWT2	SWT3
Sewer length (log)	0.793***	0.877***	0.846***	0.824***	0.885***	0.858***			
	(0)	(0)	(0)	(0)	(0)	(0)			
Pumping capacity per sewer length (log)	0.342***	0.581***	0.531***	0.352**	0.561***	0.511***			
	(0.006)	(5.6e-05)	(0.0003)	(0.0105)	(3.3e-05)	(0.0005)			
Properties per sewer length (log)	1.118***			1.076***					
	(3.4e-07)			(3e-05)					
Weighted average density—LAD from MSOA (log)		0.229***			0.257***				
		(0.00712)			(3e-05)				
Weighted average density—MSOA (log)			0.385***			0.416***			
			(0.0009)			(3e-06)			
Urban rainfall per sewer length (log)				0.104***	0.151***	0.148***			
				(0.0001)	(3e-05)	(8e-05)			
Load (log)							0.657***	0.755***	0.796***
							(6e-11)	(0)	(0)
Load treated with ammonia consent ≤ 3mg/l							0.0261		
							(0.258)		
Load treated in size bands 1 to 3 (%)							0.006***	0.006***	0.006***
							(0)	(0)	(0)
Load treated in STWs ≥ 100,000 people (%)								-0.01***	
								(0.002)	
Weighted average treatment size (log)									-0.25***
									(1.3e-06)

	SWC1	SWC2	SWC3	SWC4	SWC5	SWC6	SWT1	SWT2	SWT3
Constant	-8.116***	-6.6***	-7.6***	-7.99***	-6.4***	-7.57***	-3.76***	-4.4***	-2.99***
	(0)	(0)	(0)	(0)	(0)	(0)	(0.005)	(1e-07)	(6.1e-05)
Observations	120	120	120	120	120	120	120	120	120
Model fit	0.918	0.895	0.894	0.918	0.911	0.909	0.839	0.856	0.900
RESET	0.173	0.0770	0.0560	0.0729	0.144	0.128	0.0556	0.342	0.765
BP	6.6e-08	0	0	4.e-08	0	0	0	0	0
VIF	2.345	1.899	1.987	2.523	1.902	1.992	5.403	5.339	4.420
Pooling	0.786	0.940	0.955	0.849	0.937	0.957	0.994	0.997	0.952
Normality	0.420	0.260	0.534	0.209	0.0508	0.118	0.00056	0.0127	0.00430
Heteroscedasticity	0.272	0.0181	0.0104	0.244	0.0422	0.00825	0.892	0.276	0.406

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

The model fit of the PR24 SWC models (0.894–0.918) is lower than that of the augmented SWC models (0.920). Furthermore, the mean model fit of the PR24 SWT models (0.865) is significantly lower than that of augmented SWT model (0.900).

Table A2.3 Regression results for Ofwat's PR24 for WWNP

	WWNP1	WWNP2	WWNP3	WWNP4	WWNP5	WWNP6	WWNP7	WWNP8
Pumping capacity per sewer length (log)	0.383***	0.398***	0.371***	0.309***	0.371***	0.388***	0.359***	0.292***
	(0.00021)	(0.00031)	(0.00089)	(0.002)	(0.0005)	(0.0002)	(0.00045)	(0.00017)
Urban rainfall per sewer length (log)					0.0741**	0.0762***	0.0830***	0.0873***
					(0.0143)	(0.00879)	(0.00959)	(0.00865)
Load (log)	0.655***	0.733***	0.721***	0.722***	0.658***	0.737***	0.737***	0.727***
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Load treated with ammonia consent ≤ 3mg/l		0.0225**				0.0227**		
		(0.0313)				(0.0133)		
Load treated in size bands 1 to 3 (%)	0.0051***	0.005***	0.005***	0.005***	0.0052***	0.0051***	0.005***	0.0056***
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Load treated in STWs ≥ 100,000 people (%)			-0.00351*				-0.00411*	
			(0.0676)				(0.0623)	
Weighted average treatment size (log)				-0.098***				-0.102***
				(0.00522)				(0.00088)
Constant	-3.092***	-4.173***	-3.741***	-2.981***	-2.909***	-3.999***	-3.655***	-2.735***

	WWNP1	WWNP2	WWNP3	WWNP4	WWNP5	WWNP6	WWNP7	WWNP8
	(1.8e-05)	(3.3e-08)	(0)	(0)	(8.9e-05)	(1.8e-07)	(2.5e-09)	(0)
Observations	120	120	120	120	120	120	120	120
Model fit	0.941	0.947	0.944	0.951	0.945	0.952	0.950	0.957
RESET	0.408	0.446	0.625	0.829	0.0725	0.0445	0.00292	0.0830
BP	0	9.39e-11	0	2.91e-08	0	9.75e-07	2.86e-08	0.00403
VIF	4.244	5.459	5.339	4.442	4.348	5.460	5.378	4.604
Pooling	0.914	0.943	0.951	0.868	0.955	0.936	0.945	0.750
Normality	0.0379	0.00512	0.0415	0.0107	0.0366	0.00847	0.0594	0.00777
Heteroscedasticity	0.234	0.966	0.136	0.0980	0.0969	0.431	0.0241	0.0438

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.
Source: Oxera analysis.

The model fit of the augmented WWNP model is higher than that of any PR24 WWNP model.

A3 Bioresources models

The following section sets out Ofwat's provisional PR24 bioresources models updated with 2022/23 data.

A3.1 PR24 consultation models

The table below presents the model results from Ofwat's PR24 consultation models for bioresources total cost.

Table A3.1 Regression results for Ofwat's PR24 for BR (Part 1)

	BR1	**BR2**	**BR3**	**BR4**	**BR5**	**BR6**
Weighted average density—LAD from MSOA (log)	-0.165				-0.269	
	(0.363)				(0.211)	
Weighted average density—MSOA (log)		-0.148				-0.398
		(0.628)				(0.232)
Load treated with ammonia consent ≤ 3mg/l	0.0655*	0.0653		0.0795**		
	(0.0987)	(0.133)		(0.0253)		
Sludge produced (log)	1.204***	1.162***	1.153***	1.145***	1.061***	1.061***
	(0)	(0)	(3.91e-07)	(0)	(1.55e-08)	(4.30e-08)
Number of STWs per property (log)			0.299			
			(0.162)			
Constant	-0.884	-0.704	0.869	-1.841*	0.801	2.007
	(0.502)	(0.764)	(0.416)	(0.0265)	(0.469)	(0.331)
Observations	120	120	120	120	120	120
Model fit	0.722	0.716	0.680	0.720	0.678	0.673
RESET	0.620	0.492	0.905	0.277	0.273	0.506
BP	0.00116	0.000539	0	0.000519	0	0
VIF	3.025	3.04	3.313	2.426	2.144	2.252
Pooling	0.897	0.882	0.934	0.815	0.849	0.906
Normality	0	0	0	0	0	0
Heteroscedasticity	0	0	0.015	0	0.001	0.004

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

The table below presents the model results from Ofwat's PR24 consultation models for bioresources unit cost.

Table A3.2 Regression results for Ofwat's PR24 for BR (Part 2)

	BRU1	BRU2	BRU3	BRU4
Weighted average density—LAD from MSOA (log)		-0.216		
		(0.128)		
Weighted average density—MSOA (log)			-0.316	
			(0.146)	
Load treated with ammonia consent ≤ 3mg/l	0.0522**			
	(0.0362)			
Number of STWs per property (log)				0.179
				(0.128)
Constant	-1.042***	0.715	1.654	0.622
	(0)	(0.452)	(0.297)	(0.458)
Observations	120	120	120	120
Model fit	0.158	0.0743	0.0618	0.0719
RESET	0.662	0.0120	0.0656	0.523
BP	1.62e-06	0	0	0
VIF	1	1	1	1
Pooling	0.558	0.442	0.51	0.575
Normality	0	0	0	0
Heteroscedasticity	0.064	0.94	0.792	0.524

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

A4 Residential retail models

The following sections set out the retail model results from Oxera's assessment, and Ofwat's provisional PR24 consultation models updated with 2022/23 data.

A4.1 Augmented models

The table below presents the augmented models for RDC and ROC.

Table A4.1 Augmented models for RDC and ROC

	RDCAug1	RDCAug2	RDCAug3	ROCAug1	ROCAug1
Average bill size (£ per/household) (log)	0.752***	0.960***	0.768***		
	(0)	(0)	(0)		
Equifax—percentage of households with payment default (%)	0.0475***				
	(0.00108)				
Equifax—average number of County Court Judgements		0.455*			
		(0.0962)			
ONS—income deprivation score (interpolated) (%)			0.0708***		
			(0.000480)		
Proportion of dual-service households (%)				0.00210***	0.00222***
				(4.07e-05)	(4.62e-05)
Total number of households (log)					-0.0280
					(0.312)
Constant	-3.027***	-3.346***	-2.890***	2.730***	3.114***
	(1.45e-09)	(0)	(0)	(0)	(0)
Observations	170	170	170	170	170
RESET	0.00808	0.132	0.0182	0.179	0.141
BP	0	0	0	0	0
Model fit	0.764	0.780	0.784	0.143	0.156
VIF	1.014	1.004	1.184	1	1.707
Pooling	1	1	1	0.782	0.928
Normality	0.726	0.518	0.141	0.066	0.129
Heteroscedasticity	0.091	0.175	0.728	0.032	0.118

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

The table below presents the augmented models for RTC.

Table A4.2 Augmented for RTC

	RTCAug1	**RTCAug2**	**RTCAug3**	**RTCAug4**	**RTCAug5**	**RTCAug6**
Average bill size (£ per/household) (log)	0.609***	0.705***	0.616***	0.522***	0.607***	0.521***
	(0)	(0)	(0)	(0)	(0)	(0)
Equifax—percentage of households with payment default (%)	0.0190***			0.0199***		
	(0.00263)			(0.000754)		
Equifax—average number of County Court Judgements		0.0477			0.0613	
		(0.735)			(0.660)	
ONS—income deprivation score (interpolated) (%)			0.0226*			0.0260**
			(0.0562)			(0.0307)
Total number of households (log)	-0.0794***	-0.0892***	-0.0766***			
	(3.93e-05)	(2.87e-05)	(0.00112)			
Constant	0.576*	0.600*	0.662*	-0.0675	-0.101	0.0791
	(0.0667)	(0.0760)	(0.0626)	(0.832)	(0.766)	(0.788)
Observations	170	170	170	170	170	170
RESET	0.351	0.761	0.654	0.176	0.172	0.208
BP	0	0	0	0	0	0
Model fit	0.732	0.662	0.683	0.691	0.670	0.677
VIF	2.51	2.455	1.932	1.014	1.004	1.184
Pooling	0.995	0.953	0.995	0.991	0.932	0.935
Normality	0	0	0.003	0.001	0.002	0.002
Heteroscedasticity	0.048	0.04	0.069	0.308	0.17	0.176

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

A4.2 PR24 consultation models

The table below presents the model results from Ofwat's PR24 consultation models for RDC and ROC.

Table A4.3 Regression results for Ofwat's PR24 for RDC and ROC

	RDC1	RDC2	RDC3	ROC1	ROC1
Average bill size (£ per/household) (log)	1.105***	1.150***	1.016***		
	(0)	(0)	(0)		
Equifax—percentage of households with payment default (%)	0.0477***				
	(0.00314)				
Equifax—average number of County Court Judgements		0.729**			
		(0.0214)			
ONS—income deprivation score (interpolated) (%)			0.0722***		
			(0.00265)		
Proportion of dual-service households (%)				0.00210***	0.00222***
				(4.07e-05)	(4.62e-05)
Total number of households (log)					-0.0280
					(0.312)
Covid-19 dummy for 2019-20 (nr)	0.395***	0.376***	0.393***		
	(3.83e-06)	(5.64e-06)	(3.91e-06)		
Covid-19 dummy for 2020-21 (nr)	0.218***	0.176**	0.206***		
	(0.00472)	(0.0222)	(0.00881)		
Constant	-5.077***	-4.673***	-4.370***	2.730***	3.114***
	(5.16e-11)	(0)	(0)	(0)	(0)
Observations	170	170	170	170	170
Model fit	0.661	0.657	0.668	0.143	0.156
RESET	0.314	0.278	0.204	0.179	0.141
BP	0	0	0	0	0
VIF	1.05	1.022	1.208	1	1.707
Pooling	0.996	0.996	0.971	0.782	0.928
Normality	0	0	0	0.066	0.129
Heteroscedasticity	0	0	0	0.032	0.118

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

The model fit of Ofwat's RDC models is significantly lower than that of the augmented RDC models.

Table A4.4 Regression results for Ofwat's PR24 for RTC

	RTC1	**RTC2**	**RTC3**	**RTC4**	**RTC5**	**RTC6**
Average bill size (£ per/household) (log)	0.672***	0.722***	0.665***	0.550***	0.605***	0.539***
	(0)	(0)	(0)	(0)	(0)	(0)
Equifax—percentage of households with payment default (%)	0.0244***			0.0228***		
	(0.00170)			(0.00202)		
Equifax—average number of County Court Judgements		0.224			0.181	
		(0.155)			(0.252)	
ONS—income deprivation score (interpolated) (%)			0.0258*			0.0290*
			(0.0651)			(0.0590)
Total number of households (log)	-0.101***	-0.0997***	-0.0884***			
	(0.000180)	(0.000799)	(0.00470)			
Covid-19 dummy for 2019-20 (nr)	0.179***	0.161***	0.169***	0.176***	0.158***	0.171***
	(1.30e-09)	(4.78e-08)	(2.70e-08)	(4.91e-09)	(1.40e-07)	(4.04e-08)
Covid-19 dummy for 2020-21 (nr)	0.0609**	0.0349	0.0464*	0.0552**	0.0300	0.0455
	(0.0277)	(0.248)	(0.0913)	(0.0499)	(0.323)	(0.102)
Constant	0.370	0.511	0.489	-0.316	-0.198	-0.0850
	(0.268)	(0.135)	(0.211)	(0.495)	(0.654)	(0.838)
Observations	170	170	170	170	170	170
Model fit	0.701	0.666	0.647	0.655	0.646	0.640
RESET	0.646	0.529	0.443	0.427	0.0736	0.329
BP	0	0	0	0	0	0
VIF	2.548	2.467	1.934	1.05	1.022	1.208
Pooling	1	1	1	1	1	1
Normality	0.031	0.034	0.052	0.011	0.021	0.027
Heteroscedasticity	0.026	0.02	0.033	0.168	0.094	0.095

Note: *, **, and *** show statistical confidence at the 90%, 95% and 99% levels, respectively.

Source: Oxera analysis.

Five of the six Ofwat RTC models have a significantly lower model fit than their smoothed equivalents (the augmented models). One Ofwat RTC model remains largely at the same model fit as its augmented equivalent.

oxera

Contact

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

oxera.com