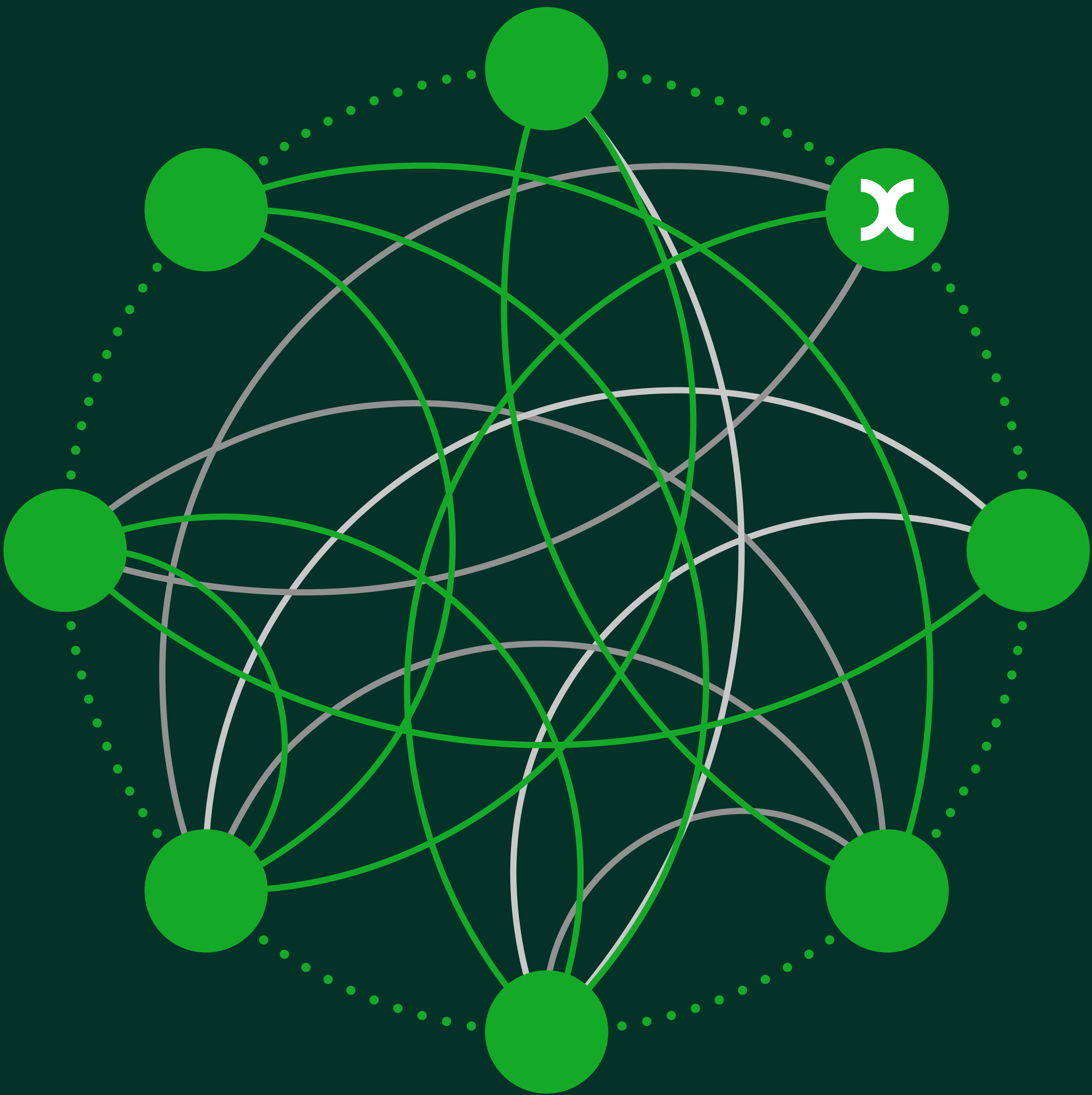
Cost adjustment claims

—

Prepared for Yorkshire Water Services

27 August 2024



Contents

Figures and Tables

Cost adjustment claims

Introduction

Ofwat has acknowledged that some cost drivers for specific companies may not be adequately reflected in its base cost models.[1] To address this, it allowed companies to submit cost adjustment claims (CACs) alongside their business plans, so that Ofwat can make suitable post-modelling adjustments where justified.

As part of its business plan submission, Yorkshire Water Services (YWS) commissioned Oxera to provide top-down evidence to support two of its proposed CACs, relating to combined sewers and phosphorus removal (P-removal) activity.[2] YWS submitted a further three CACs relating to infrastructure maintenance activity, non-infrastructure maintenance activity and meter renewal activity. In the PR24 Draft Determination (DD), Ofwat rejected YWS's CACs in relation to combined sewers and non-infrastructure maintenance activity,[3] but provided industry-wide adjustments for the following issues.[4]

- Increased mains replacement (i.e. infrastructure maintenance) activity.
- Increased meter renewals activity.
- Increased P-removal activity.
- Sustained increases in energy prices.
- Reductions in greenhouse gas (GHG) emissions.

YWS has commissioned Oxera to review Ofwat's rejection of YWS's CACs and its approach to determining the industry-wide post-modelling adjustments.

Mains replacement

Ofwat noted that it is concerned by the low rates of mains replacement in recent years, and argues that this is likely to lead to a deterioration in asset health in the near future.[5] To address this, Ofwat has introduced a

[1] Ofwat's cost models have been derived through econometric modelling. Ofwat (2024), 'PR24 draft determinations: Expenditure allowances - Base cost modelling decision appendix', July.
[2] See Oxera (2023), 'An assessment of Yorkshire Water's cost adjustment claims', September.
[3] Ofwat (2024), 'Base cost adjustment claim feeder model – Yorkshire Water', https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-YKY_Cost-adjustment-claims.xlsx, last accessed 15 August.
[4] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances', July, section 2.2.
[5] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances', pp. 30–37.

price control deliverable (PCD), whereby companies are subject to financial penalties if they do not replace necessary water mains in a timely fashion in AMP8. As part of this PCD, Ofwat has allowed some companies additional expenditure to accommodate mains replacement activity beyond the level that is implicitly funded through Ofwat's cost models. Specifically, this adjustment consists of four components.

1. **Target rate**: the target rate of mains replacement activity (for YWS, this is determined to be 0.66% p.a.).
2. **Implicitly funded rate**: the implicitly funded rate of mains replacement activity, which Ofwat assumes is the historically delivered mains replacement activity over the modelling period in 2012–23 (0.3% p.a.).
3. **Under-delivered rate**: the extent to which companies' assets have deteriorated, which Ofwat considers to be an underspend of historical allowances (for YWS, this is determined to be 0.14% p.a.).
4. **Efficient unit cost**: the median unit cost of mains replacement activity across a sub-sample of the industry.

The post-modelling adjustment is defined as the difference between the target rate (1) and the sum of the implicitly funded and under-delivered rate (2+3), multiplied by the efficient unit cost (4).

There are several issues with how Ofwat has applied the adjustment.

First, Ofwat is incorrect to assume that this implicitly funded rate is based on the average activity in the modelling period. Ofwat benchmarks cost performance over the last five years of outturn data (now, 2020–24), such that the models implicitly fund companies for the mains replacement activity in that period, or 0.15% p.a. In essence, this is due to the fact that mains replacement activity is an omitted variable as it is not included in Ofwat's models. In deriving the implicitly funded activity level, Ofwat assumes that the omitted variable is uncorrelated with the cost drivers included in its models. Otherwise, the implicitly funded activity would depend on the level of correlation and be company-specific. As such, the cost impact of mains replacement activity feeds into the constant in the regression. However, the constant in the regression is adjusted based on the performance of companies in the benchmarking period, such that the benchmarking period (not the modelling period) is the determinant of what is implicitly funded.[6]

Second, Ofwat's assessment of YWS's 'deterioration' in asset health is misleading. Ofwat examines the increase in the proportion of YWS's network that is in grades 4 and 5 (the 'poorest' network condition) between PR09 and PR24, and concludes that YWS assets have deteriorated as a result of 'under-delivery'. However, the proportion of YWS's network that is in grade 1 (the 'healthiest' grade) has increased from 36.8% in PR09 to 46.2% in PR24. Moreover, the rate of mains burst has declined materially from c. 269 in 2012 to c. 219 in 2023, suggesting that the quality of YWS's asset base has improved over time. To state that YWS has under-delivered on its past commitments based on one measure of asset health while excluding other relevant measures is misleading, and results in a material reduction in the adjustment for YWS. Given that there is no clear evidence that YWS's assets have deteriorated, or that the supposed deterioration has resulted in a consumer detriment,[7] we consider that the under-delivery adjustment should be removed.

Third, the median unit cost of mains replacement is based on a sample of 11 companies, and there is a wide range of unit costs across these 11 companies (the highest unit cost is c. three times larger than the lowest unit cost). Such a wide variation in unit costs could indicate that the unit costs are not measured consistently across companies, or that there are company-specific drivers of unit costs that are not captured in Ofwat's modelling.[8] We consider that Ofwat should investigate the data in more detail to assess whether a median unit cost should be applied, or whether a company-specific unit cost would be more appropriate.

The table below shows YWS's adjustment for mains replacement retaining a target replacement rate of 0.66% and correcting for the errors in Ofwat's calculation.

[6] This is discussed in more detail in Appendix A1.
[7] Although the evidence suggests that YWS's assets have not deteriorated over time and that any potential deterioration has not materially affected consumer outcomes, we note that the current rate of mains replacement activity may still be unsustainably low. That is, the low rates of mains replacement activity could lead to poorer consumer outcomes in future, such that more replacement activity needs to be funded in future (i.e. as part of this adjustment). This is consistent with Ofwat's view of the industry as a whole, where it argues that the low replacement rates have not led to a deterioration in asset health but may do so in the future. See Ofwat (2024), 'PR24 draft determinations: Expenditure allowances', July, pp. 30–31.
[8] The unit cost data comes from a mix of company-specific queries, leakage enhancement queries and resilience enhancement. Given that the sources of the unit cost information differ across the sample, it is likely that there are some reporting inconsistencies.

Table 1 Mains replacement

	Ofwat DD	Oxera
Target rate	0.66%	0.66%
Implicitly funded rate	0.30%	0.15%
Under-delivery rate	0.14%	0%
Efficient unit cost (£/m)	292	292[1]
Total adjustment (£m)	106	248

Note: [1] We consider that there may be issues with the data that are distorting Ofwat's assessment of the efficient unit cost of mains replacement activity, but we are not in a position to comment as to what the efficient cost should be.
Source: Oxera analysis.

The total mains replacement adjustment for YWS increases from £106m to £248m under this approach, which is a c. 134% increase. This suggests that Ofwat's current methodology is materially underfunding YWS for mains replacement activity, which would increase the risk that YWS is unable to deliver on its maintenance programme to the detriment of consumers and the environment.

Meter renewals

At a high level, Ofwat's approach to meter renewals is similar to its approach to mains replacement. That is, Ofwat acknowledges that several companies are expecting a step-change in meter renewals in AMP8 that are not fully captured in the cost assessment models, such that an adjustment to their allowances is required. To calculate the adjustment, Ofwat follows the same four steps as outlined in the mains replacement section. Therefore, the issues with the mains replacement adjustment outlined above are also relevant for the meter renewals adjustment, specifically in relation to the implicitly funded rate of meter renewals and the efficient unit cost.

Correcting for these issues results in an increase in YWS's adjustment from £164m to £166m.

Energy price adjustment

Ofwat argues that the historical cost models do not fully fund companies for the recent increase in energy prices.[9] Specifically, Ofwat argues that the econometric cost models will fund companies for the average energy price observed in the modelling period (2012–23), which is lower than the current price that companies are facing. However, Ofwat also considers that energy prices will reduce over AMP8. Therefore, Ofwat constructs the energy price adjustment in a two-step approach, as follows.

1. **Uplift**: Ofwat calculates an uplift to companies' expenditure based on the difference between the energy price in 2023 (the final year of the modelling period) and the average energy price in the modelling period (2012–23). Ofwat uses the DESNZ index to measure energy prices, which is an index of the energy prices facing industrial consumers and therefore accounts for the hedging strategies of industrial consumers. This results in a positive adjustment to companies' allowances.
2. **Real Price Effect (RPE)**: Ofwat uses Bloomberg forecasts of the spot price of energy to predict energy prices in AMP8. As the forecasts show that energy prices are expected to decline in AMP8, this results in a negative adjustment to companies' allowances.

The principal concern with Ofwat's analysis is that there is a disconnect between Ofwat's approach to determining the uplift and Ofwat's approach to forecasting energy prices. Specifically, the index used to construct the uplift accounts for industrial hedging strategies and may therefore be expected to reasonably track the prices that companies face, while the index used to construct the RPE relates to the spot price of energy. The disconnect between the two is apparent when Ofwat's modelling is updated with the latest outturn data: the forecasts at the DD predicted that energy prices would materially decline in 2023/24 (by c. 34%), yet the DESNZ index has actually increased in that period (by c. 13%).

Ofwat should update its modelling to reflect the latest outturn data. As shown in the table below, this would result in a significant increase in YWS's energy price uplift.

[9] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances', pp. 44–46.

Table 2 Energy price uplift

	Water resources	Water network plus	Wastewater network plus	Bioresources	Total
Ofwat's DD (£m)	-0.8	-7.8	-11.1	1.3	-18.3
Oxera's approach (£m)	4.4	45.4	64.7	-6.0	108.5
Impact (£m)	5.2	53.2	75.8	-7.3	126.8

Source: Oxera analysis.

In addition, we consider that the energy RPE should be based on an energy price index that accounts for hedging strategies, in the same way as the uplift is calculated. One option is to use companies' forecasts of energy prices to estimate the RPE. While this data may be endogenous, consumers are protected from any overestimation risk via the true-up mechanism.

Phosphorus removal

In its PR24 business plan, YWS submitted a CAC relating to increased operating expenditure (OPEX) resulting from additional and more complex P-removal activities due to tightening environmental regulation.[10] Ofwat has accepted the need for adjustments relating to increased P-removal activity, given that its models do not explicitly account for P-removal and that the tightening of P-consent levels is a relatively 'new' activity such that it is unlikely that companies are implicitly funded through the models. Therefore, Ofwat has applied an industry-wide post-modelling adjustment to account for the increased expenditure requirements in relation to this issue.[11]

Specifically, Ofwat has developed an econometric cost model (estimated at the sewerage treatment works (STW) level) to estimate the relationship between P-consent level and OPEX, and uses this model to predict companies' incremental costs associated with increased P-removal activity. Ofwat's model is fairly simple: OPEX (in £m) is modelled as a function of population equivalent (PE) and P-consent level. At a high level, this is broadly aligned with one of the approaches that we developed for YWS to estimate a CAC for the business plan

[10] See Oxera (2023), 'An assessment of Yorkshire Water's cost adjustment claims', September, section 3.
[11] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances', July, p. 40.

submission. However, we consider that Ofwat's model specification is flawed, for two reasons.

First, Ofwat estimates the model in terms of levels rather than logarithms. This imposes an unintuitive restriction on the relationship between P-consent level and OPEX, in that a decrease in P-consent has the same monetary impact on STWs' OPEX (c. £23k per mg/l reduction) for all STWs, regardless of their size. This is likely to be inconsistent with operational intuition—an STW that treats more load (i.e. is larger) will need to incur more additional costs (in monetary terms) as a result of tighter P-consent levels, given that more chemicals and energy would be required to treat more load. Ofwat's modelled functional form is also not supported by the RESET test, which is Ofwat's preferred test for model misspecification.

To correct for this, we consider that it would be appropriate to estimate the model is logarithms rather than in levels. Under this specification, a tightening in the P-consent level would have a proportionate impact on costs depending on STW size. We note that this functional form relationship is better aligned with Ofwat's approach to assessing expenditure in other areas, including its base expenditure models and some enhancement models.

Second, we consider that the models should allow for a more flexible relationship between PE and costs by accounting for the step-change in costs that occur at particularly tight P-consent levels. Amending the model in this way results in a statistically and economically superior model specification that has statistically significant coefficients and passes the RESET test.

Moreover, Ofwat has applied an upper-quartile (UQ) benchmark when estimating the adjustment. The range of efficiency scores in Ofwat's model is particularly wide (42–192%), which suggests that there is significant uncertainty in the models as they are sparsely specified. Therefore, we consider that a median benchmark is more appropriate,[12] and this is aligned with Ofwat's assessment of enhancement expenditure where there is a similar degree of uncertainty.[13]

[12] We note that the range of efficiency scores is similarly wide in our better-specified models, suggesting that there is still significant uncertainty in the modelling.
[13] For example, see Ofwat (2024), 'PR24 draft determinations: Expenditure allowances - Enhancement cost modelling appendix', July, pp. 2–7.

The table below shows YWS's adjustment under these modelling corrections.

Table 3 Phosphorus removal

	UQ benchmark	Median benchmark
Ofwat DD models	87	119
Oxera models	116	165

Source: Oxera analysis.

Combined sewers

Due to the additional maintenance and repair costs associated with combined sewers, as they are more prone to sewer flooding and increased pressure than separate sewers, we consider combined sewers to be an operationally relevant driver of costs. During periods of heavy rainfall, combined sewers can exceed their design capacity, leading to blockages, partial collapses and flooding incidents. To mitigate such occurrences, additional infrastructure, such as storage tanks, may be necessary to store and divert excess flows, which increases the complexity and costs associated with the sewerage network.

YWS has the second-highest proportion of combined sewers in the industry. This means that the company is materially affected by the exclusion of combined sewers in Ofwat's PR24 models. In the DD, Ofwat did not include combined sewers, arguing that (i) the driver does not have a clear engineering rationale; and that (ii) it could drive perverse incentives as a result of endogeneity.[14]

In the PR24 modelling consultation, 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).[15] We do not consider that urban rainfall captures similar cost pressures to combined sewers. First, the two drivers are not strongly correlated with each other (the correlation coefficient is c. 0.26), which suggests that these variables are likely to capture

[14] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances – Base cost modelling decision appendix', July, p. 45.
[15] Ofwat (2023), 'Econometric base cost models for PR24', April, p. 45.

different effects. This lack of correlation is unsurprising, given that the relative abundance of combined sewers is driven largely by managerial decisions made before privatisation,[16] while urban rainfall is driven by a combination of population density and climate. That is, while density and climate may have had an influence on managerial decisions in relation to the network, the two drivers are not conceptually related to each other,[17] other than by the fact that both affect costs, as evidenced by the low correlation between the two.

Second, the two cost drivers perform well when included in the same model on the current dataset—the cost drivers are both statistically significant and are (directionally) aligned with expectations, and the inclusion of both drivers improves model fit. This provides empirical evidence that the two drivers could capture different costs and can therefore be included in the same model.

As the combined sewer systems were also installed prior to privatisation in 1989, we consider that the proportion of combined sewers is largely outside management control, particularly in the short term, given the difficulty and costly nature of upgrading or changing the structure of such systems. In principle, the combined sewers variable may be considered endogenous in the sense that it relates to the physical assets that companies deploy, in the same way that nearly all of the cost drivers included in Ofwat's wholesale models are endogenous.[18] To that extent, the combined sewers variable can be deemed to be exogenous in the short term, in the same way as other measures of physical assets that are included in Ofwat's model are considered exogenous. Other regulators also use physical assets as a cost driver. For example, Ofgem incorporates modern equivalent asset value (MEAV) as a driver in its TOTEX (total expenditure) modelling, and the CMA endorsed models that controlled for physical assets as cost drivers at the PR19 redetermination.

If companies had material control over the length of combined sewers, there would be strong incentives for them to reduce this, given that: (i) combined sewers are associated with higher costs (as supported by the econometric cost modelling); and (ii) companies have not historically received uplifts or additional allowances to compensate for having

[16] Indeed, we understand that the practice of installing combined sewers was largely discontinued after World War II.
[17] The two drivers could only be conceptually related insofar as historical managerial decisions were driven by the population density and climate at the time. However, we are unaware of evidence to suggest that this is the case.
[18] The only truly exogenous drivers in Ofwat's wholesale models are connected properties, population density and urban rainfall.

more combined sewers. However, we note that the proportion of combined sewers across the industry has been static across the modelling period (2012–24), with just a 0.66 percentage point decrease in the industry-average proportion of combined sewers, of which YWS has had one of the greatest reductions in proportion by 0.86 percentage points. Despite the incentive to reduce the length of combined sewers, it is clear that companies have been unable to do this to any material extent, which indicates that the risk of endogeneity is potentially limited.[19]

Finally, Ofwat refers to Dŵr Cymru (WSH) as an example of a company that has a high proportion of combined sewers and urban rainfall while performing well on internal sewer flooding (ISF). We do not consider that an isolated example of one company's circumstances is a robust reason to conclude that combined sewers is not an operationally relevant driver of costs. We note that there is a statistically significant *positive* relationship between combined sewers and ISF, as well as urban rainfall and ISF. That is, higher proportions of both combined sewers and urban rainfall are associated with greater ISF incidents. This suggests that both drivers affect performance separately.[20]

Furthermore, while WSH has the highest urban rainfall per sewer length, has a relatively high proportion of combined sewers and performs comparatively well on ISF, it performs less well on external sewer flooding (ESF). If total sewer flooding (TSF, the sum of ISF and ESF) is considered, the relationship between combined sewers and flooding incidents is even stronger, and WSH no longer appears to be an outlier. A company's strong performance on ISF and comparatively weak performance on ESF may be driven by allocation issues or by differing management focus on flooding incidents. In either case, relying on one data point in one measure of flooding incidents will provide an incomplete assessment.

We estimate the impact of combined sewers on YWS's efficient cost prediction across AMP8 by taking the difference in predictions between Ofwat's DD models and where combined sewers are included in its

[19] In principle, it might be possible to test 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.
[20] See Oxera (2024), 'Addressing the disconnect between cost and outcomes', August.

sewage collection (SWC) and network plus (WWNP) models.[21] This results in an estimated net CAC value of c. £43m for YWS across AMP8.

In addition to this CAC, we also explored a performance adjustment claim in relation to ISF, given that the presence of combined sewers affects companies' abilities to mitigate ISF. To derive an adjustment, we have estimated econometric models of incidence of ISF. To isolate the impact of combined sewers, we have then compared YWS's predicted incidence of ISF based on its forecast level of combined sewers abundance to its hypothetical predicted incidence of ISF if it had the industry average level of combined sewers abundance. Our results indicate that the adjustment is c. 0.77 incidents per 10,000 connections per year in AMP8, which would be the performance adjustment. That is, YWS has 0.77 more ISF incidents per 10,000 connections than the industry average purely as a result of combined sewers.[22] We note that there are other drivers of ISF, including population density and urban rainfall, that also indicate that YWS should have more ISF incidents than the rest of the sector, although these are not directly related to the performance adjustment.

GHG emissions

Ofwat has made additional allowances for base expenditure to account for companies' efforts to decarbonise the network. As with the mains replacement and meter renewals adjustment, the Net Zero adjustment comes with a PCD to ensure that companies actually deliver on GHG emissions in a timely fashion.

In constructing the adjustment, Ofwat has assessed the unit cost of individual schemes proposed by companies, defined as pounds per GHG reduced. The median unit cost across these schemes is then applied to the GHG reduction target (2.47% reduction, which is based on the median reduction in GHG emissions) to derive the adjustment. There are several issues with the way in which this adjustment is calculated.

First, it is apparent from Ofwat's modelling that different schemes to reduce GHG emissions are associated with different unit costs. Taking a

[21] To be consistent with Ofwat's DD, we apply a UQ benchmark on the efficient cost predictions. We provide an assessment of the appropriate benchmark in a separate report: Oxera (2024), 'Base cost modelling at PR24', August.
[22] We note that there are other drivers of ISF, including population density and urban rainfall, that also indicate that YWS should have more ISF incidents than the rest of the sector, although these are not directly related to the performance adjustment. Using econometric models to inform service targets is discussed in more detail in Oxera (2024), 'Addressing the disconnect between cost and outcomes', August.

median across all schemes will penalise companies for taking on high-cost schemes. While this may make sense if all schemes were available to all companies and the industry were starting 'from scratch', it may penalise companies that have already taken steps to reduce carbon emissions. For example, replacing the vehicle fleet with electric vehicles (EVs) appears to be associated with lower unit costs (median unit cost c. £758 per tCO_2e) than installing heat pumps (median unit cost c. £2,288 per tCO_2e). However, if a company has already replaced a significant proportion of its fleet with EVs (the lower-cost option), it may be forced to adopt schemes with higher unit costs in order to meet the targets.

Second, the median unit cost is derived from 17 schemes across five companies. This is a particularly small sample size given the heterogeneous nature of the schemes and is distorted by the presence of multiple schemes for Thames Water. Indeed, if the average unit cost is constructed at the company level rather than at the scheme level, the unit cost (and therefore the adjustment) increases by c. 25%. In the absence of additional data that can enable a more robust benchmarking of costs at the scheme level, estimating the benchmark at the company level may be more reflective of the mix of schemes that companies have proposed.

A more robust framework could involve benchmarking companies' unit costs at a scheme level, such that (for example) schemes relating to heat pumps are benchmarked only against other schemes relating to heat pumps. This may require more data, given that this type of scheme-level information does not appear to be available for all companies (or, at least, the data was not deemed to be sufficiently robust to have fed into the calculation of the post-modelling adjustment).

Summary

Ofwat's post-modelling adjustments at the DD are a means to provide additional funding for companies to deliver outcomes for consumers in AMP8 that are incremental to base cost allowance. However, Ofwat has made a series of errors in each adjustment that materially underestimate the required adjustment for YWS. The table below compares the adjustment that YWS received at the DD relative to what the adjustment would be under a more robust methodology.

Table 4 Summary of cost adjustment claims

	Ofwat DD	Improved methodology
Mains replacement (£m)	106	248
Meter renewals (£m)	164	166
Energy uplift (£m)	-18	109
Phosphorus removal (£m)	87	165
Combined sewers (£m)	0	43
Combined sewers (internal sewer flooding incidents per 10,000km)	0	0.77
GHG emissions (£m)	5	6
Total adjustment (£m)	344	737

Source: Oxera analysis.

Our analysis suggests that Ofwat's adjustments for YWS should be more than double what it has proposed at the DD, which indicates that the DD materially underfunds YWS on account of these issues.

A1 Why the benchmarking period?

A critical issue with Ofwat's assessment of 'what base buys' that is common across several of Ofwat's post-modelling adjustments is that the implicitly funded level of activity is estimated as the industry-average activity throughout the modelling period (2012–23).

In essence, Ofwat's post-modelling adjustments amount to an 'omitted variable' problem—there is some driver of expenditure (such as mains replacement activity) that Ofwat acknowledges influences costs, yet is not accounted for in the econometric model. Therefore, we can assess the implicit allowance by examining how an omitted variable influences the cost models and subsequently a company's efficient expenditure.

First, we assume that the omitted driver is uncorrelated with the cost drivers included in the models. In this case, the omitted driver can be treated as a random, weakly positive variable. Suppose that Ofwat's models are otherwise unbiased and that there is only one omitted factor. The true cost function is:

$$\ln(Cost_{it}) = \beta_0 + \beta_1 * \ln(Cost\ driver_{it}) + \gamma * (Omitted\ driver_{it}) + \varepsilon_{it}$$

Where:

- $Cost_{it}$ is the observed cost of company i at time t;
- $Cost\ driver_{it}$ is the observed cost driver of company i at time t;
- $Omitted\ driver_{it}$ is the observed omitted driver of company i at time t;
- ε_{it} is statistical noise for company i at time t.

However, Ofwat estimates the following regression.

$$\ln(\widehat{Cost_{it}}) = \widehat{\beta_0} + \widehat{\beta_1} * \ln(Cost\ driver_{it})$$

Where the 'hat' indicates that these are estimated values of the true parameters. Given that we have assumed that the omitted driver is uncorrelated with the cost drivers in the model, the estimated $\widehat{\beta_1}$ is unbiased. However, the estimated $\widehat{\beta_0}$ **is** biased, as it contains the cost impact of the average omitted activity over the modelling period—i.e. the implicitly funded level of the omitted activity. In this stylised case, it would be broadly appropriate to determine the implicitly funded level of activity as the industry-average activity over the period (i.e. Ofwat's approach at the DD).

However, this stylised case is unlikely to accurately reflect the current context. For example, the stylised case assumes that the cost drivers are uncorrelated with the omitted factor. If, instead, there is a strong correlation between the cost drivers and the omitted factor, then the estimated coefficient on the cost driver (i.e. $\widehat{\beta_1}$) would be biased. Specifically, the estimated coefficient would capture some of the cost impact of the omitted driver, such that the implicitly funded level of activity would differ by company depending on the value of that cost driver. Nonetheless, assuming that the omitted factor is uncorrelated with the cost drivers may be an appropriate and proportionate simplifying assumption in some cases. For example, we found that meter renewal activity is not correlated with the cost drivers included in Ofwat's models, such that it may be simpler and more appropriate to assume that the driver is uncorrelated.

More importantly, while the constant in Ofwat's regression analysis is estimated using the modelling period (2012–23), the constant that is used to set allowances is adjusted and determined by the benchmarking period (2019–23). This is because Ofwat adjusts allowances based on the performance of companies in the last five years, such that Ofwat's estimated efficient cost function is not necessarily the pure output from the regression. Instead, while the coefficients of the cost drivers are indeed the pure output from the regression, the constant is adjusted based on the performance in the last five years. Given that the value of the constant is informed entirely by companies' performance in the last five years, the implicitly funded level of activity is also the industry average over the last five years (again, assuming that the omitted activity is uncorrelated with the cost drivers).

Given that Ofwat corrects to the UQ benchmark, the degree to which the omitted activity is implicitly funded is technically driven by the average activity of the UQ company. However, we do not consider that it would be appropriate to determine what is implicitly funded on the basis of one company, given that:

- strictly speaking, companies would be funded for all of the omitted factors related to the UQ company, not just the omitted activity in question;
- the company may also have undertaken an exceptionally low or high level of the omitted activity as a direct decision of management, given prior flexibility on what companies were able to direct funding to;
- relying specifically on the benchmark company may result in unjustified volatility if there are any changes to the model specification or the benchmark stringency.

For these reasons, we consider that it is appropriate to assess the implicitly funded level of activity on the basis of the industry-average performance during the benchmark period, unless there is sufficient evidence that the omitted activity is strongly correlated with the cost drivers included in the models.

A2 Mains replacement

A2.1 Background and Ofwat's proposed approach

Ofwat argues that the health of the water industry's network assets has generally improved over time, pointing to the reduced rates of mains bursts and improvements to more granular measures of asset health. Nonetheless, Ofwat states that it is concerned with the low rates of mains replacement across the industry in recent years (c. 0.1% p.a.) and that companies are not undertaking enough renewal activity to keep up with asset deterioration. Therefore, Ofwat has developed a PCD to require companies to replace assets that are currently in poor condition, whereby companies receive financial penalties if they do not replace the target volume of assets in a timely fashion. Alongside this PCD, Ofwat has provided some companies with an uplift to their base expenditure allowances if their required maintenance activity is above what is implicitly funded through the models.

Specifically, this adjustment consists of four components.

First, Ofwat sets the **target rate** of mains replacement activity that companies are required to deliver as part of the PCD. For most companies, this is the rate of mains replacement activity that is implicitly funded through the models (discussed in more detail below). However, for companies whose asset health is deemed to be 'below average', Ofwat requires them to deliver more than what is implicitly funded. For most of the affected companies, this rate is set at 0.43% p.a. For YWS, the rate is set at 0.66% p.a., as Ofwat has agreed with YWS's justification for a higher rate.

Second, Ofwat determines what is **implicitly funded** through the models. Ofwat assumes that companies are implicitly funded to deliver the average mains replacement activity observed in the modelling period (2012–23), or 0.3% p.a. As noted in Appendix A1 this is incorrect: if mains replacement activity were uncorrelated with the cost drivers in Ofwat's models, the implicitly funded rate of mains replacement activity would be the average rate in the benchmarking period (2019–23 at the time of the DD, currently 2020–24).

Third, Ofwat assesses the volume of activity that (in its view) companies have previously funded but have not delivered—i.e. the '**under-delivered rate**'. Ofwat's assessment assumes that companies have historically been funded to maintain the health of their assets, where 'asset health' in this context refers to the proportion of assets in condition grades 4 and 5. For companies that have seen an improvement in asset health according to this measure from PR09 to PR24, Ofwat applies no under-delivery adjustment. For companies that have seen a deterioration in asset health according to this measure, Ofwat estimates that under-delivery rate as the average annual deterioration in asset health from PR09 to PR24.

Fourth, Ofwat constructs an efficient **unit cost** for mains replacement activity, calculated as the median unit cost in companies' business plans or responses to queries.

Our concerns with Ofwat's approach to constructing this adjustment are outlined in more detail below and in Appendix A1.

A2.2 Improvements to Ofwat's adjustment

A2.2.1 Under-delivery adjustment

Conceptually, we consider that it is inappropriate to apply an 'under-delivery adjustment' to the implicitly funded level of mains replacement activity. Ofwat's concern stems from its assumption that companies have been historically funded to deliver a certain level of mains replacement. However, this is not the case: companies are provided with base (or TOTEX) allowances, and it is up to them to allocate funding to individual activities in a way that is aligned with consumers' interests and the regulator's objectives. If, for example, companies have been set particularly stretching cost and performance commitments, they may have to reduce non-critical maintenance expenditure in order to achieve performance within their allowances. Similarly, if companies face exogenous shocks during a regulatory period (for example, macroeconomic shocks such as an increase in energy prices), they may have to re-prioritise activities to keep expenditure within their allowances (or minimise the degree of overspend).

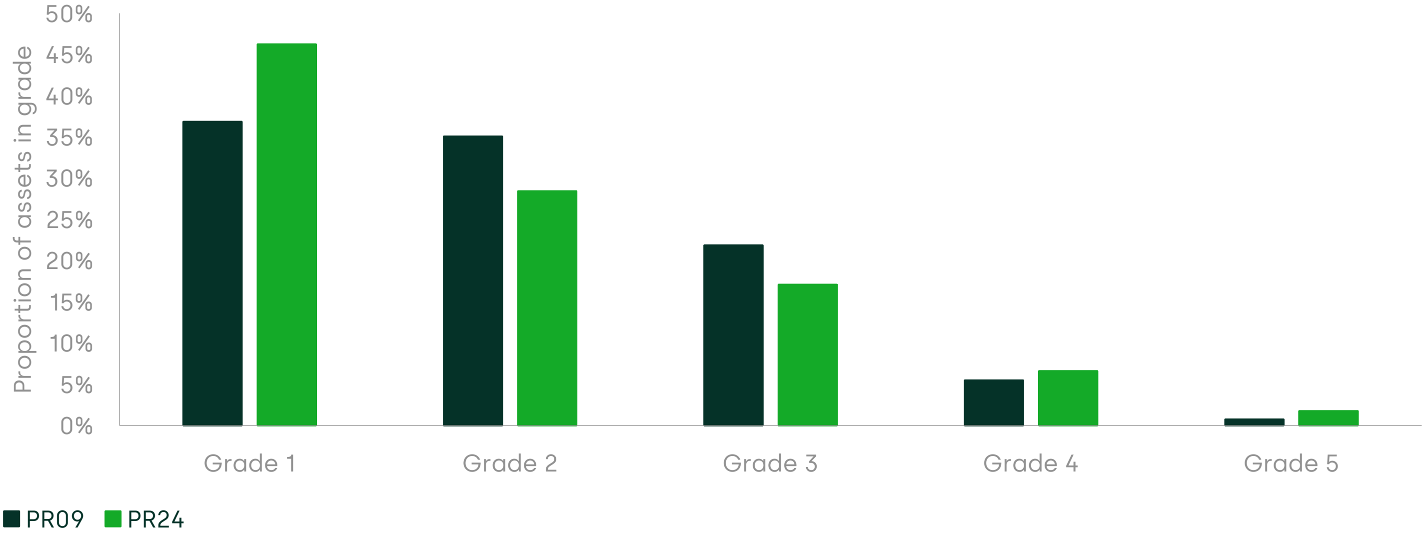
Of course, Ofwat may be concerned that companies have under-delivered if:

- they have not undertaken critical maintenance activity and subsequently outperformed their cost allowances, thereby increasing returns for shareholders at the expense of consumers; and/or
- Ofwat had provided discrete funding packages for individual activities, with associated PCDs.

However, we note that neither of these were the case in AMP7.

In addition to these concerns, we note that Ofwat's conclusion that the health of YWS's assets has deteriorated is based on one, partial, measure of asset quality. A more balanced review of the evidence suggests that YWS's assets have not deteriorated over time. For example, the figure below compares the proportion of mains in each condition grade in PR09 with PR24.

Figure A2.1 YWS's asset health



Source: Oxera analysis of Ofwat (2024), 'Mains renewal cost adjustment model'.

The figure shows that YWS has materially increased the proportion of its network that is in grade 1 (the 'healthiest' grade), from c. 37% in PR09 to c. 46% in PR24. If we define 'healthy' assets as being assets within the top two grades (equivalent to Ofwat defining 'poor' assets as those in the bottom two grades), YWS's asset health has improved from 72% in PR09 to 75% in PR24. YWS's average condition grade has also improved from c. 1.98 in PR09 to 1.89 at PR24. As such, the weight of evidence suggests that YWS's assets are currently of a higher quality than they were at PR09.

As an alternative to grading the condition of assets, we have also explored how YWS has performed on the performance commitments that are most relevant to mains asset health. YWS's level of service has typically improved materially over the period for which comparable data is available. For example:

- YWS's rate of mains bursts has reduced from c. 269 incidents per 10,000km in 2012 to c. 219 incidents per 10,000km in 2023, a reduction of c. 19%;[23]
- YWS's supply interruptions have reduced from c. 19 minutes in 2012 to c. 9.5 minutes in 2023, a reduction of c. 50%;
- YWS's leakage has reduced from c. 323 Ml/d in 2018 to c. 283 Ml/d in 2023, a reduction of c. 12%.

That is, YWS's level of service on these measures has materially improved over time, which casts doubt on the conclusion that the quality of YWS's assets has materially deteriorated.

Given that there is no strong evidence that YWS's assets have deteriorated, and that the weight of evidence suggests that their quality has improved, it is inappropriate to impose an under-delivery penalty on YWS.

A2.2.2 Unit cost comparisons

Ofwat sets the efficient unit cost as the median unit cost for mains replacement activity proposed in companies' business plans and in response to queries from Ofwat. Ofwat excludes the unit cost for Thames Water (TMS) when calculating the median, arguing that there are London-specific drivers that make TMS's unit costs not comparable with the rest of the industry. The table below summarises the unit costs for each company.

[23] Ofwat examined the rate of mains bursts across the industry, and found that asset health had improved. See Ofwat (2024), 'PR24 draft determinations: Expenditure allowances', July, p. 30.

Table A2.1 Mains replacement unit costs

Company	Unit cost (£/m)	Source	Ofwat calculation?
TMS	1,458	TMS18	No
YKY	336	YKY45	No
WSH	420	WSH62	Median of submitted costs
WSX	350	WSX09	No
NES	274	NES35	No
ANH	273	Enh - resilience	
SRN	661	Enh - resilience	
SWB	292	Leakage query	Average cost based on length and expenditure
BRL	280	Leakage query	Average cost based on length and expenditure
SVE	310	Leakage query	Average cost based on length and expenditure
NWT	218	Leakage query	Average cost based on length and expenditure
SEW	274	Leakage query	Average cost based on length and expenditure

Source: Ofwat (2024), 'Mains renewal cost adjustment model', https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-Mains-renewals-adjustments.xlsx, last accessed 19 August 2024.

The table shows that there is a wide range of unit costs across the industry, even if TMS is excluded from the calculation. The average unit cost and the spread of unit costs appear to vary depending on the source:

- queries from Ofwat: £274–£1,458 per meter with an average of £568 per meter (excl. TMS, this is £274–£420 and £345 per meter, respectively);
- enhancement resilience: £273–£661 per meter with an average of £467 per meter;
- leakage query: £218–£310 per meter with an average of £275 per meter.

Given the wide range in unit costs across the industry, and the fact that the average and spread of unit costs differ across different sources of information, we consider that it is likely that these unit costs are not capturing solely differences in efficiency across companies. Instead, it is likely that the spread in unit costs is driven by data errors, inconsistencies in reporting, or heterogeneity in operating environments.

We consider that Ofwat should interrogate the data in more detail to assess whether the unit costs reported by companies are capturing the same activities, and explore whether there are regional cost pressures that would warrant a company-specific unit cost for mains replacement activity.

A3 Phosphorus removal

A3.1 Background and Ofwat's proposed approach

YWS faces tightening environmental regulations that require additional and more complex P-removal activities and thus increased operating expenditure (OPEX). Since P-removal activities are not adequately captured by the cost drivers in Ofwat's DD models for wastewater, Ofwat proposes a post-modelling adjustment to account for the additional cost associated with these activities.

Ofwat's proposed post-modelling adjustment uses STW-level data to estimate the relationship between P-consent level and OPEX. The model's dependent variable is OPEX in levels, which is regressed on the cost drivers population equivalent (PE) served and assumed P-consent. Ofwat has determined that some observations are 'outliers' based on the Cook's distance outlier metric, and has removed these observations from the sample when estimating its model. The results are reported for estimating the model with pooled ordinary least squares (OLS) and random effects, but only the results from the random effects regression are then used to predict companies' incremental costs associated with increased P-removal activity. Table A3.1 shows the regression results.

Table A3.1 Results from Ofwat's proposed Draft Determination models for the post-modelling adjustment relating to P-removal

	PM1	PM2
PE served	0.003***	0.003***
P-consent (assumed)	-0.022***	-0.023***
Constant	0.071***	0.074***
Dependent variable	OPEX	OPEX
Estimation method	Pooled OLS	Random effects
Model fit	0.606	0.606
RESET p-value	0.000	0.000

Note: *** reflects statistical significance at the 1% level. P-consent (assumed) reflects the P-consent level in mg/l. Where the STW does not have a permit, this variable takes the value 5. PE served is in 10,000s. OPEX is in £m.
Source: Ofwat.

The coefficients are statistically significant and their signs are aligned with expectations. That is, higher PE served and tighter P-consent levels, which require more complex treatment procedures, are associated with higher costs. The results indicate that a 1mg/l decrease in the P-consent level is associated with an increase in incremental OPEX of £23k, regardless of the size of the STW or the amount of load it has to treat. The model fails the RESET test for misspecification with a P-value close to 0, suggesting that there are non-linear and/or interaction effects between the independent variables.[24]

Ofwat then uses the modelled coefficients to predict costs at the STW level for all STWs, including the outlier STWs that are not included in the modelling. These predictions are then aggregated at the company level and efficiency scores are calculated to determine final allowances based on a UQ benchmark.

We consider that Ofwat's proposed modelling is flawed, for the following reasons.

Economies of scale: Ofwat estimates its model in levels rather than in logs. The latter specification is used for nearly all other expenditure assessments, including the wastewater base cost modelling. While the coefficients in Ofwat's model are directionally intuitive, the economic meaning of the coefficients is difficult to justify. In particular, the model specification implies that the returns to scale vary depending on the P-consent level. This is illustrated by the equations below. Equation I shows the model with the parameter estimates from regression PM2:[25]

$$\text{(I)} \qquad OPEX = 0.074 + 0.003 * PE\ served + (-0.023) * PE\text{-}consent$$

Now we assume that STW A has a PE consent level of 1mg/l and STW B has a P-consent level of 4mg/l. This yields the following equations.

$$\text{(II)} \qquad OPEX_{STW\ A} = 0.051 + 0.003 * PE\ served$$

$$\text{(III)} \qquad OPEX_{STW\ B} = -0.018 + 0.003 * PE\ served$$

Since the P-consent level is linear, it is effectively added to the constant. The model would predict negative OPEX for small values of PE served for STW B. Now we assume that both STWs have a value of PE served of 10.

$$\text{(IV)} \qquad OPEX_{STW\ A} = 0.081$$

[24] The null hypothesis of the RESET test is that the model is correctly specified or, more precisely, that higher-order exponents of the independent variables do not have explanatory power once the independent variables are controlled for.

[25] All values are rounded to the third digit after the decimal separator.

(V) $OPEX_{STW\ B} = 0.012$

Now we assume a doubling of the PE served from 10 to 20 for both STWs, which yields the following results.

(VI) $OPEX_{STW\ A} = 0.111$
(VII) $OPEX_{STW\ B} = 0.042$

This shows that a doubling of scale implies a 37% increase in predicted OPEX for STW A and a 250% increase in predicted OPEX for STW B, even with the same size (PE served) as a starting point. The difference in returns to scale is caused solely by the difference in P-consent levels. This appears to be difficult to justify operationally, and we are not aware of any evidence to support this assumption.

Interaction between P-consent and scale: the model assumes that the additional OPEX associated with lower P-consent levels is independent of the PE served. This is not in line with operational insights, which suggest that this OPEX is a variable cost, since a larger amount of load treated requires more energy and/or chemicals to be treated at a given P-consent level.

Non-linear impact of P-consent activities: the model assumes that the impact of tighter P-consent levels is linear. This contrasts with operational insights, which suggest that particularly low levels of P-consent are associated with disproportionately high OPEX. Ofwat acknowledges this non-linearity in its post-modelling adjustment related to enhancement P-removal, for which it includes a dummy reflecting P-consent close to the technically achievable limit (TAL) in one of the econometric models used to set allowances.[26]

Benchmark selection: Ofwat has selected a UQ benchmark (estimated at the company level) to adjust the predicted costs, despite the estimated efficiency scores ranging from 0.42 (ANH) to 1.92 (NES). The wide range of efficiency scores suggests that there is significant uncertainty in this modelling, and that an average or median benchmark is more appropriate.

A3.2 Our proposed alternative approach

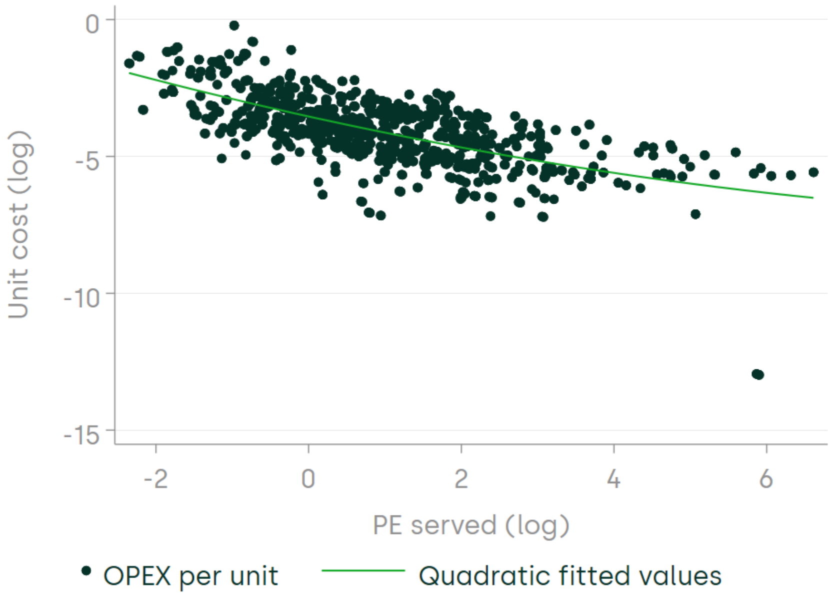
To address the shortcomings of Ofwat's DD models, we have developed alternative models for the P-removal post-modelling adjustment. These

[26] The TAL is assumed to be 0.25mg/l.

alternative models are based on the same data as that used by Ofwat and use the same approach to excluding outliers.

To adequately model economies of scale, we investigate unit costs relative to STW size. Figure A3.1 shows a scatter plot of the unit cost by STW against the size of the STW, as measured by the load received.[27]

Figure A3.1 Economies of scale at the STW level



Note: Unit cost (log) reflects the natural logarithm of OPEX in £m/PE served in 10,000s. PE served (log) reflects the natural logarithm of PE served in 10,000s.
Source: Oxera analysis based on Ofwat data.

The scatter plot suggests that the unit cost decreases with the size of the plant, implying positive economies of scale. The green line reflects the fitted values of a regression of unit costs on PE served, as well as squared PE served. It indicates that economies of scale are decreasing with the size of the plant. We therefore include a linear and a quadratic term of PE served as control variables in our models.

We model unit costs as a function of P-consent levels, controlling for economies of scale. The table below shows two alternative models: one controlling for the P-consent level only, and one controlling for the P-

[27] For ease of interpretation, we have modelled the logarithm of the unit costs as a function of the STW size.

consent level and a dummy variable indicating a P-consent level below 0.5mg/l.

Table A3.2 Results from regressions of unit costs on P-consent and economies of scale variables

	PM_ALT1	**PM_ALT2**
P-consent (assumed)	-0.272***	-0.111***
P-consent below 0.5mg/l		0.407***
Log(PE served)	-0.687***	-0.716***
Squared log(PE served)	0.057***	0.060***
Constant	-3.271***	-3.565***
Dependent variable	Log(OPEX per PE served)	Log(OPEX per PE served)
Sample size	1183	1187
Model fit	0.493	0.536
RESET test p-value	0.702	0.742
Estimation method	Random effects	Random effects

Note: *** reflects statistical significance at the 1% level. Standard errors are clustered at the company level. P-consent (assumed) reflects the P-consent level in mg/l; when the STW does not have a permit, this variable takes the value 5. PE served is in 10,000s. OPEX is in £m.
Source: Ofwat.

All cost drivers are associated with the expected sign and are statistically significant at the 1% level. The results from regression PM_ALT1 indicate that a 1mg/l decrease in the (assumed) P-consent level is associated with a c. 31% increase in unit costs.[28] The results from regression PM_ALT2 indicate that 1mg/l in the assumed P-consent level is associated with a c. 12% increase in unit costs, while also indicating that P-consent levels below 0.5mg/l are associated with an additional increase in unit costs by c. 50%. The model fit is high for a unit cost model.

We consider that these models constitute an improvement on Ofwat's model specification, for the following reasons.

[28] This is calculated from the coefficient β_1 for P-consent (assumed) x_2 as follows: $\Delta \log(unit\ cost) = \beta_2 \Delta x_2 \Leftrightarrow \Delta unit\ cost = \exp(\beta_2 \Delta x_2) - 1$.

- The models account more flexibly for economies of scale than Ofwat's proposed models. The modelling suggests that an increase in scale is associated with a decrease in costs (i.e. economies of scale), but that the extent of the economies of scale diminishes as STW-size increases. In contrast to Ofwat's proposed models, our alternative models pass the RESET test.
- The models allow for tightening P-consent levels to have a proportionate rather than an absolute impact on OPEX, which is aligned with operational expectations.
- The models account for the fact that particularly tight P-removal levels below 0.5mg/l are associated with disproportionately higher OPEX (as shown by the positive and significant coefficient on the P-consent below 0.5mg/l dummy variable).
- All cost drivers are statistically significant at the 1% level, and their sign is consistent with operational expectations.

However, even with these improvements, the range of efficiency scores remains quite wide, which justifies a median benchmark.[29]

A3.3 Predicted efficient costs

We use the results from models PM_ALT1 and PM_ALT2 to predict incremental OPEX at the STW level for all STWs in 2026, including those excluded from the modelling. This is in line with the approach taken by Ofwat.

We then aggregate the predictions at the company level and assume constant expenditure over AMP8. Table A3.3 shows the results for a median and a UQ benchmark below.

[29] The range of efficiency scores is 1.45 in Ofwat's models and 1.90 in the alternative models (post triangulation). This difference is due mainly to Northumbrian Water (NES), which is assessed to be the least efficient company under both approaches. When excluding NES, the range of efficiency scores is 1.33 in Ofwat's models and 1.40 in the alternative models.

Table A3.3 YWS's allowance relating to the P-removal post-modelling adjustment

Model	Benchmark	Allowance (£m)
Ofwat Draft Determinations	Upper quartile	86.7
	Median	119.0
Alternative models	Upper quartile	115.6
	Median	164.7

Source: Oxera analysis based on Ofwat data.

The table shows that the alternative models suggest that the efficient cost of YWS's expected incremental P-removal activity in AMP8 is in the range of £115.6m–£164.7m. YWS's submitted costs of £108.7m are below this range.

A4 Combined sewers

The sewerage network is designed to handle both surface water (such as rainfall) and wastewater (or 'foul' water). This network can consist of either separate sewers for surface water and foul water or a single network of combined sewers. Historically, before privatisation, combined sewers were often installed because they generally required less space, given that only one pipe was needed instead of two separate ones for surface water and foul water.

However, during periods of heavy rainfall, combined sewers can exceed their design capacity, leading to blockages, partial collapses and flooding incidents. To prevent such occurrences, additional infrastructure, such as storage tanks, may be necessary to store and divert excess flows, which increases the complexity and costs associated with the sewerage network. Due to these additional costs, it has become more common for companies to install separate networks for surface water and foul water.

At PR19, Oxera proposed models that accounted for combined sewers as a cost driver. Several companies also submitted models that included combined sewers as a cost driver in the PR24 modelling consultation. This indicates that there is at least some support from the industry for considering combined sewers in the assessment of companies' costs.

However, in the PR24 DDs Ofwat did not include combined sewers, arguing that (i) the driver did not have a clear engineering rationale; and (ii) it could drive perverse incentives as a result of endogeneity.[30] In the PR24 modelling consultation, 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 companies' control).[31]

A4.1 Combined sewers and urban rain cover different cost pressures

Ofwat's suggestion that including urban rainfall as a cost driver has a similar impact to including the percentage of combined sewers appears to be based on the following logic.

[30] Ofwat (2024), 'PR24 draft determinations: Expenditure allowances – Base cost modelling decision appendix', July, p. 45.
[31] Ofwat (2023), 'Econometric base cost models for PR24', April, p. 45.

- Combined sewers are more prone to flooding, so the costs associated with them are typically linked to managing sewer flooding.
- Urban rainfall is also intended to capture costs relating to sewer flooding.
- Since urban rainfall already accounts for a characteristic that increases sewer flooding, there is no need for an additional cost driver that captures similar costs, such as combined sewers.

However, this reasoning is flawed. The fact that urban rainfall increases sewer flooding does not imply that combined sewers do not also increase flooding. The two cost drivers are not intrinsically linked and cannot be treated as proxies or substitutes. The relative abundance of combined sewers is driven largely by managerial decisions made before privatisation, while urban rainfall is driven by a combination of population density and weather/climate.[32]

Moreover, all the cost drivers in Ofwat's models are 'high-level' and may capture various characteristics. For instance, using mains or sewer length as a scale variable may partially reflect additional costs of operating in sparsely populated areas—where a company in a sparser region is likely to have greater mains length than a similarly sized company in a denser area. However, it would be incorrect to assume that population density is unnecessary in the cost models simply because the scale variable partially captures the impact of density or sparsity.

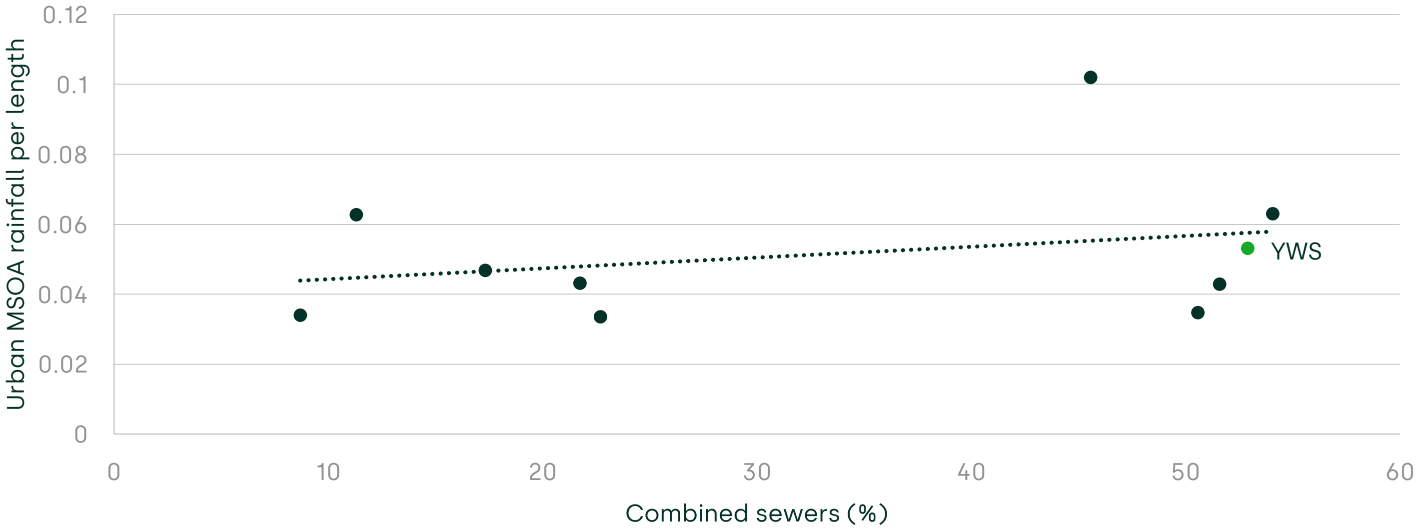
Ofwat's current approach assumes that cost drivers should be grouped into categories based on how they influence costs (e.g. scale, complexity and topography), and that only one cost driver should be selected from each category. In this case, Ofwat has grouped combined sewers and urban rainfall into the same category ('costs associated with flooding'). However, these drivers could just as easily be placed in different categories (e.g. 'climate and weather' and 'network complexity'), which would not justify excluding combined sewers from a model that includes urban rainfall. This categorisation issue highlights a fundamental problem with this modelling approach, which is furthermore not fully aligned with regulatory precedent. For instance, at PR19, Ofwat controlled for both population density and STW size in its

[32] Ofwat, 'Water sector overview', https://www.ofwat.gov.uk/regulated-companies/ofwat-industry-overview/, accessed 15 August 2024.

bioresources (BR) models, even though both were intended to capture the cost impact of STW-level economies of scale.

For urban rainfall to sufficiently capture the additional costs associated with combined sewers, there would have to be a strong, ideally perfect, correlation between the two cost drivers. Figure A4.1 shows the relationship between companies' average urban rainfall and average percentage of combined sewers over the period from 2012 and 2024.

Figure A4.1 Relationship between urban rainfall and combined sewers (2012–23)



Note: YWS is highlighted in green. MSOA, middle layer super output areas.
Source: Oxera analysis.

While the figure indicates that there is some correlation between urban rainfall and combined sewers, this relationship is relatively weak with a low correlation coefficient of 0.26. In other words, urban rainfall only captures a small share of the variability in combined sewers across companies.

In the case of YWS, its share of combined sewers is higher than its level of urban rainfall would suggest (i.e. it is below the regression line). Therefore, failing to account for combined sewers leads to biased outcomes for YWS, even if one assumes that the cost-impact of combined sewers is already captured by the urban rainfall driver (as noted above, this is not the case).

Furthermore, it is possible to directly test whether combined sewers and urban rainfall both drive internal sewer flooding (ISF) using econometric models. At this stage, we consider that the following factors may drive ISF.

- Population density and sparsity—it may be easier (i.e. quicker) to resolve issues on the network in some regions than in others due to population density and/or sparsity.
- Asset health—assets that are in better condition are less likely to fault and therefore less likely to result in leakage.
- Urban rainfall—companies that operate in regions that experience heavy rainfall are more prone to flooding.
- Network configuration—combined sewers are more prone to sewer flooding than other types of asset.

As a consequence, we use the following independent variables to model ISF: weighted average MSOA to LAD and weighted average MSOA to control for density; urban rainfall per length to control for urban rainfall; and the share of combined sewers to control for network configuration. To facilitate an intuitive interpretation and allow for non-linear effects, we use the natural logarithm of all variables other than the share of combined sewers. Table A4.1 shows the regression results.

Table A4.1 Results from regressions of incidences of internal sewer flooding on combined sewer, urban rainfall and density

	ISF1	ISF2
Combined sewer (percentage)	0.0179*	0.0207**
Urban MSOA rainfall per length (log)	0.405*	0.408*
Weighted average density MSOA to LAD (log)	0.250	
Weighted average density MSOA (log)		0.635*
Constant	-0.658	-3.961
Observations	60	60
Model fit	0.1685	0.2719

Note: ** and * reflect statistical significance at the 5% and 1% level respectively.
Standard errors are clustered at the company level.
Source: Oxera analysis based on Ofwat data.

All independent variables have the expected sign and are statistically significant, with the exception of weighted average density MSOA to LAD in ISF1. The coefficients indicate that a one percentage point increase in the share of combined sewers is associated with a c. 2% increase in the number of sewer flooding incidents per 10,000 connections. The fact that both the combined sewers and the urban rainfall cost drivers are statistically significant in the same model indicates that they are separate drivers of internal sewer flooding.

A4.2 Network configuration with respect to combined sewers cannot be materially changed in the short run

Ofwat uses 'asset-based' cost drivers across its modelling suite, where companies have some control over the driver in the long run but not in the short run, including:

- the length of the water network in Ofwat's treated water distribution (TWD) models;
- the length of the sewer network in Ofwat's SWC and WWNP models;
- the size of treatment works in Ofwat's sewage water treatment (SWT), WWNP and BR models;
- the number of booster pumping stations in Ofwat's TWD and wholesale water (WW) models.

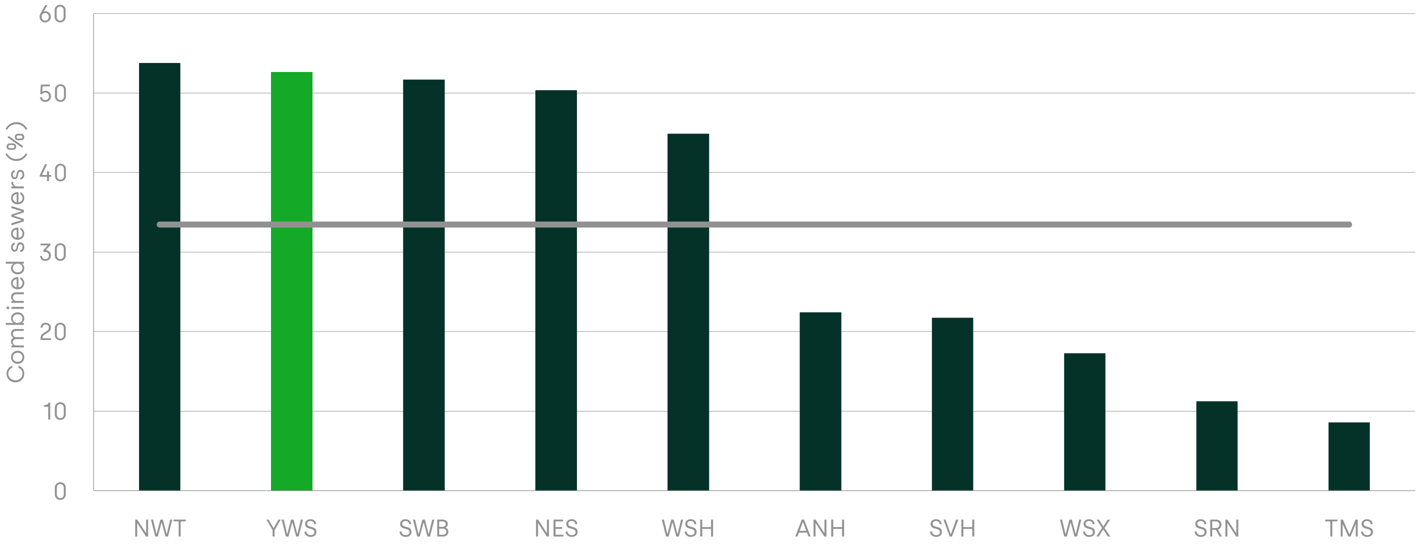
We consider that Ofwat's argument—that companies may be incentivised to invest in combined sewers to receive higher cost allowances—is unrealistic. The proportion of combined sewers has not materially changed for any company over the last 12 years, indicating that the extent to which companies have any substantial control over this variable is limited. Across the industry, companies have reduced the proportion of combined sewers by only 0.05 to 1.90 percentage points, with an industry-average reduction of 0.66 percentage points.

In the current context, combined sewers are associated with higher costs, yet these high costs are not reflected when setting cost allowances. If combined sewers were indeed endogenous in the short run, companies would have had strong incentives to reduce their number of combined sewers, for example by replacing them with separate surface water and foul water sewers, in order to perform better in the cost assessment models and achieve higher returns.

A4.3 Why an adjustment for YWS is required

YWS has a particularly high percentage of combined sewers. Figure A4.2 shows the proportion of combined sewers for each wastewater company as well as the industry average.

Figure A4.2 Percentage of combined sewers (2020–24)



Note: The grey line reflects the industry average.
Source: Oxera analysis.

The figure shows that YWS has the second-highest proportion of combined sewers in the industry (c. 53%), behind United Utilities (NWT). This is c. 19 percentage points above the industry average. Therefore, the observation that combined sewers are not accounted for in the cost models is likely to lead to a downward-biased estimate of YWS's efficiency.

A4.4 Empirical analysis: wastewater cost assessment models including a combined sewers cost driver

The most direct approach for estimating the cost impact of combined sewers is to include this cost driver explicitly in the PR24 cost models. The table below shows how Ofwat's wastewater models perform when the percentage of combined sewers is included as an additional cost driver.

Table A4.2 Wastewater models (SWC and WWNP) with combined sewers cost driver

	SWC1	SWC2	SWC3	WWNP1	WWNP2
Sewer length (log)	0.835***	0.944***	0.914***		
Pumping capacity per sewer length (log)	0.411**	0.720***	0.661***	0.500***	0.391***
Density (log)	1.322***				
Urban MSOA rainfall per sewer length	0.0764**	0.127**	0.125**	0.0417**	0.0586*
Combined sewers (percentage)	0.0025*	0.0049*	0.0053**	0.0028***	0.0022*
Weighted average density—LAD to MSOA (log)		0.331***			
Weighted average density—MSOA (log)			0.546***		
Load (log)				0.805***	0.759***
Proportion of load treated in size bands 1–3 (%)				0.0252***	
Proportion of load treated with ammonia consents <3mg/l (%)				0.00492***	0.00554***
Weighted average treatment plant size (log)					-0.0838***
Constant	-0.0838***	-0.0838***	-0.0838***	-0.0838***	-0.0838***
Observations	130	130	130	130	130
Model fit	0.912	0.914	0.911	0.951	0.950
RESET (p-value)	0.0000	0.0126	0.00980	0.00853	0.00340

Note: ***, ** and * reflect statistical significance at the 1%, 5% and 10% level, respectively. Standard errors are clustered at the company level.
Source: Oxera analysis based on Ofwat data.

The coefficient on combined sewers is positive and statistically significant in all models. The coefficient associated with urban MSOA rainfall per sewer length is also positive and statistically significant in all models. In addition, the inclusion of combined sewers leads to an improvement in model fit relative to Ofwat's models, ranging between 0.2 percentage points in SWC1 and 1.4 percentage points in SWC3.

The fact that the two cost drivers reflecting urban rainfall and combined sewers perform well when included in the same model on the current dataset provides empirical evidence that the two drivers capture different cost pressures.

Table A4.3 shows how YWS's allowance is affected by the omission of combined sewers in the cost assessment models. We have applied a UQ benchmark to the predicted costs in each suite of models. Therefore, the cost predictions and CAC can be considered efficient.

Table A4.3 Combined sewers CAC value (£m)

Model	Allowance (£m)
Ofwat DD models	1,805.9
DD models with combined sewers	1,849.2
Difference	43.2

Note: The allowances presented exclude any RPEs and frontier shift.
Source: Oxera analysis based on Ofwat data.

The table shows that the net CAC value relating to combined sewers is c. £43.2m over AMP8. This is above Ofwat's 1% materiality threshold for WWNP CACs. In line with Ofwat's modelling guidelines, the implicit allowance is the efficient cost prediction under the PR24 DD cost models, and the gross CAC value is the efficient prediction under models that control for combined sewers.

A4.5 Performance adjustment relating to internal sewer flooding

In addition to incurring additional costs as a result of combined sewers (which would necessitate a CAC), YWS incurs more ISF incidents (which would necessitate an adjustment to the performance commitment level, PCL). The relationship between ISF and combined sewers is demonstrated empirically in the section above.

To quantify this performance adjustment, we use the results from the econometric models reported in Table A4.1. As explained above, these models provide estimates of the impact of combined sewers on the incidence of ISF.

We predict YWS's incidence of ISF over AMP8 based on its forecast values for the share of combined sewers and other performance drivers. We then predict YWS's incidence of ISF for the hypothetical scenario in which it has the industry-average share of combined sewers. This corresponds to the implicitly funded level of ISF incidence. All predictions are transformed into levels and the results from the two models are triangulated. We compare the predicted values for YWS with the forecast share of combined sewers with those predicted for YWS with the industry share of combined sewers. The results are presented in Table A4.4.

Table A4.4 Combined sewers performance adjustment

	Incidence of ISF
Forecast share of combined sewers	2.49
Industry-average share of combined sewers	1.72
Difference	0.77

Note: Incidence of ISF in incidents per 10,000 connections.
Source: Oxera analysis based on Ofwat data.

The predicted incidence of ISF of 2.49 based on YWS's forecasts of combined sewers is materially higher than the predicted incidence of ISF of 1.72 based on the industry-average share of combined sewers. The difference amounts to 0.77 incidents per 10,000 connections per year over AMP8.

oxera

Contact

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

oxera.com