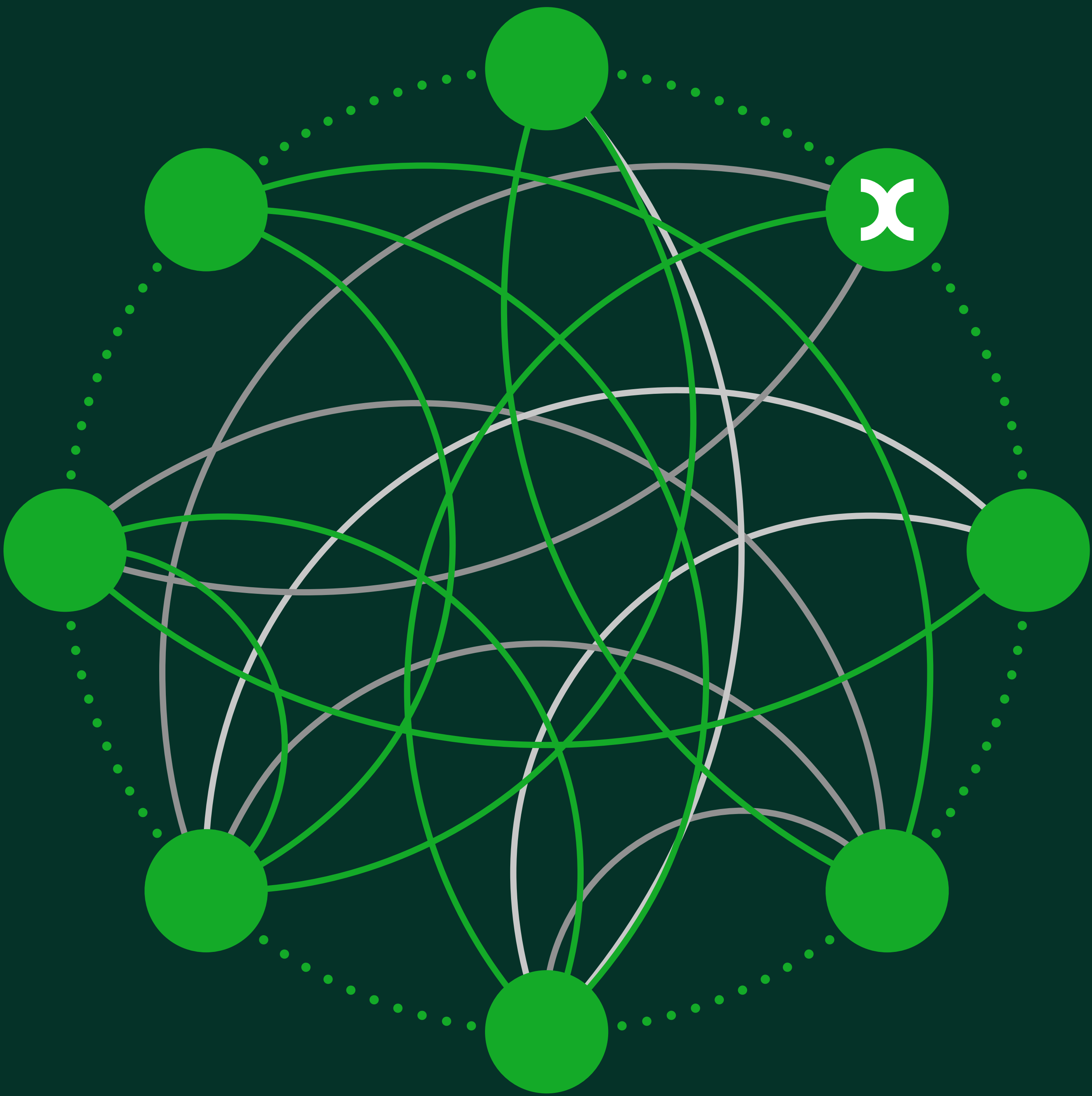
Addressing the disconnect between cost and outcomes

—

Prepared for Yorkshire Water Services

27 August 2024



Contents

Figures and Tables

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

Addressing the disconnect between cost and outcomes

Introduction

Economic theory dictates that an efficient company can improve service quality only through additional resources or productivity improvements.[1] For this reason, it is important for regulators to consider the overall 'stretch' on the cost envelope and performance commitments (PCs) when assessing whether their regulatory determinations are achievable. As part of this assessment, the regulator should assess the level of performance that could be reasonably delivered through the base expenditure allowances and, if a more challenging level of performance is required (e.g. due to strong customer preferences or government obligations), the additional funds that would be needed to achieve these targets.

At PR19, Ofwat's assessment of costs was largely disconnected from its assessment of performance commitment levels (PCLs). Specifically, expenditure allowances on base and enhancement activities were determined through econometric modelling, unit cost comparisons, or deep or shallow dives,[2] while the PCLs were generally determined using one or more of the following approaches:[3]

- considering the upper-quartile (UQ) performance forecast in companies' business plans;
- assuming a specific rate of improvement relative to the PR14 PCL;
- considering what would be required to meet full compliance with statutory measures;
- considering the UQ rate of improvement proposed in companies' business plans;

[1] Indeed, when Ofwat outlines how companies can expect to improve service, it provides examples of increased investment (i.e. additional resources) and improved management practices (i.e. productivity improvements). For example, in its discussion on storm overflows, Ofwat states: '[w]e consider that companies can deliver improvements in storm overflow spills by preventing or removing blockages, undertaking investment to ensure existing permits are met and maintaining assets well'. See Ofwat (2024), 'PR24 draft determinations: Delivering outcomes for customers and the environment', July, p. 47.
[2] Ofwat (2019), 'PR19 final determinations: Securing cost efficiency technical appendix', December.
[3] Ofwat (2019), 'PR19 final determinations: Delivering outcomes for customers policy appendix', December.

- calculating the average of a company's 'best-performing' outturn years.

That is, Ofwat did not undertake a robust assessment of what companies could reasonably deliver through their expenditure allowances, such that there was insufficient evidence that the overall stretch on cost and outcomes was achievable. Indeed, four companies appealed the PR19 determination to the Competition and Markets Authority (CMA) in part as a result of this issue. While the CMA allowed some disputing companies additional expenditure to meet stretching PCLs,[4] it did not itself establish a robust methodology for integrating the assessment of cost and service within the time frame of the redeterminations.

Shortly after the CMA redetermination, Ofwat stated that it would undertake a more robust assessment of 'what base buys' with respect to service performance:[5]

Our ambition for PR24 is to build on our PR19 approach to setting cost allowances and performance levels by drawing a more explicit link between cost allowances and the service levels we set (i.e. the cost-service relationship).

Despite the stated intention, Ofwat did not propose a concrete approach for doing so in its PR24 methodology. In the PR24 draft determination (DD), Ofwat has still not undertaken a thorough assessment of the level of service that could reasonably be delivered through base expenditure. Instead, it has (by and large) set PCLs independently of its assessment of expenditure, as at PR19.

Yorkshire Water (YWS) has commissioned Oxera to review the disconnect between Ofwat's approach to cost assessment and its approach to performance assessment, and assess whether this disconnect leads to an unjustified productivity challenge and thereby asymmetric operational risks.

Ofwat's approach at the DD

[4] For example, Anglian Water and Yorkshire Water received higher allowances in order to meet the stretching leakage PCL. See 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. 8.205.
[5] Ofwat (2021), 'Assessing base costs at PR24', December, p. 66.

For most PCLs, Ofwat has generally assumed that the PR19 PCL for 2025 has already been funded, such that the PR19 PCL forms the baseline ('year 0') target relative to which companies may be expected to make further improvements.[6] However, as noted above, there was a disconnect between Ofwat's assessment of costs and its assessment of PCLs at PR19, such that there is actually no evidence that the PR19 PCLs have already been funded. Ofwat's assumptions may be justified if companies had managed to meet or exceed their targets on costs and service in AMP7. However, not a single company has met both its cost and service targets in AMP7, with the vast majority of the sector underperforming on both.[7] That is, there was no evidence at the time of PR19 that the PR19 PCLs were achievable, and no new evidence has come to light that suggests that the PR19 PCLs were achievable. Ofwat should present some evidence that its year 0 PCL target is achievable, rather than assuming that it has already been funded.

Ofwat uses various sources of evidence to support or validate its proposed level of stretch beyond the year 0 target, most of which do not involve an assessment of what is funded through Ofwat's cost allowances. These sources of evidence, and our concerns regarding why they do not account for the link between cost and service, are outlined below.[8]

- **Extrapolating historical performance**—companies are not funded for their rate of improvement in terms of service, but (at most) for the level of service that they have achieved. For example, if a company increases the monitoring of its network (i.e. increasing costs) to reduce leakage, it may be able to maintain this improved level of leakage if it continues to monitor its network at the heightened levels. However, it cannot be expected to further reduce leakage without again increasing the monitoring of its network at additional cost. Note that Ofwat does not apply this logic to its assessment of post-modelling adjustments, where it assumes that companies are funded for the average *level* of activity and not the *trend or rate*.[9]
- **Companies' proposed improvements in their business plans**—companies do not propose PCLs in isolation, but as part of a package with both other PCLs and cost allowances. Ofwat has

[6] See Ofwat (2024), 'PR24 draft determinations: Delivering outcomes for customers and the environment', July, section 3.
[7] This is discussed in more detail in appendix A1.
[8] Ofwat (2024), 'PR24 draft determinations: Delivering outcomes for customers and the environment', July, section 8.
[9] This is discussed in more detail in appendix A3.

made material challenges to companies' TOTEX requests, such that it cannot be assumed that companies can deliver the stretching PCLs that they proposed within Ofwat's allowance. In any case, Ofwat's comparisons across companies do not account for regional factors that affect the ability of companies to achieve stretching PCLs (a stretching PCL may be relatively easier/cheaper to achieve in one region than in another).[10]

- **Companies' compliance with statutory obligations**—it cannot be assumed that companies are funded to deliver stretching PCLs simply because these relate to long-term government objectives or statutory obligations, particularly if they involve improvements to companies' current levels of service.
- **An average of companies' 'best-performing' years**—a company's performance on some service measures, such as the number of mains bursts and sewer collapses, in a given year will be driven (in part) by stochastic events (such as the weather or asset failure). To cherry-pick companies' performance in the best-performing years would not account for these stochastic events, meaning that the overall target may not be achievable.

That is, Ofwat's approach to assessing PCLs is largely divorced from its assessment of expenditure, such that it cannot be assumed that the PCLs are implicitly or explicitly funded through Ofwat's cost allowances. Therefore, a robust methodology is required to assess what is funded through the models.

Note that Ofwat (and, indeed, companies) may wish to set PCLs that go beyond what is funded through Ofwat's cost allowances. For example, if some PCs are particularly important for consumers or government objectives, it would be appropriate for companies to deliver more material improvements. However, companies may only be able to achieve such stretching PCLs if they are adequately funded to do so.

An improved methodology

We have considered four approaches of varying sophistication to assess the extent to which service is implicitly funded by the models, as outlined below.

Approach 1: industry-average performance

[10] This is discussed in more detail in appendix A2.

As a starting point, it may be appropriate to determine what is implicitly funded through Ofwat's base expenditure models by examining the average performance of the companies over the benchmark period (i.e. by considering the last five years of outturn data). This is broadly consistent with how Ofwat assesses the extent to which mains replacement activity, meter renewal activity and energy prices are implicitly funded in the models for its post-modelling adjustments.[11] As with Ofwat's approach to determining the post-modelling adjustments, this approach assumes that service performance is *uncorrelated* with the cost drivers included in the models and that there are no other drivers of service performance. While this is a fairly strong (and testable) assumption, the approach is comparatively simple (both conceptually and computationally) and is consistent with Ofwat's approach elsewhere. It is therefore easily implementable for the final determination (FD).

Approach 2: predicted performance—cost drivers

For some PCs, it is probable that companies' performance is correlated with some of the cost drivers in the models, such as cost drivers relating to population density (included in both wholesale water and wholesale wastewater models) and urban rainfall (included in wholesale wastewater models). Therefore, the estimated coefficients in the cost models will capture, in part, some of the costs associated with achieving different levels of service performance. That is, companies would be implicitly funded for a different level of service depending on the relationship between service quality and the cost drivers included in the models. In this case, the correlation between service performance and the cost drivers would need to be accounted for when determining what is implicitly funded through the models.

Ofwat makes the same conceptual argument when deciding to accept or reject companies' cost adjustment claims (CACs), as in the following examples.

- Regional wages: Ofwat has rejected claims relating to regional wages, in part arguing that the cost impact of regional wages is captured by the density drivers in the cost models.[12] That is,

[11] The principal distinction between Approach 1 and Ofwat's approach to assessing the implicit allowance is that Approach 1 focuses on the benchmark period (2019–23) whereas Ofwat focuses on the modelling period (2012–23). The reasoning behind this is outlined in appendix A4.
[12] For example, see Ofwat (2024), 'Base cost adjustment claim feeder model – Affinity Water', found here https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-AFW_Cost-adjustment-claims.xlsx, last accessed 6 August 2024.

companies are funded for a particular level of regional wages depending on their level of density.
- Economies of scale at the water treatment work (WTW) level: Ofwat has rejected claims relating to WTW-level economies of scale, in part arguing that the cost impact of WTW-level economies of scale is implicitly captured by the density drivers in the cost models.[13] That is, companies are funded for a particular average WTW size depending on their level of density.
- Liming and bioresources: Ofwat has rejected a claim relating to the additional costs associated with particular treatment technologies, arguing that treatment technology is correlated with treatment work size.[14] That is, companies are implicitly funded for particular treatment technologies depending on the size of their treatment works.

In each of these cases, Ofwat has argued that the costs associated with the 'omitted factor' are implicitly captured through correlations between the omitted factor and the cost drivers included in the models. However, in rejecting the claims, Ofwat does not present evidence of *the extent to which* the cost drivers capture these omitted factors.

In the context of PCs, we consider that the extent to which the cost drivers capture differences in service (and therefore implicitly fund different levels of service) could be estimated through the following procedure.

1. For each relevant cost model, regress the service measure against the cost driver included in that model.
2. Use these models to predict the level of service for each company.
3. Triangulate the predicted level of service across the models, following Ofwat's triangulation approach.

This approach is fairly simple to implement under the current framework.[15] Additional complexities may arise if Ofwat were to make

[13] For example, see Ofwat (2024), 'Base cost adjustment claim feeder model – South East Water', found here https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-SEW_Cost-adjustment-claims.xlsx, last accessed 6 August 2024.
[14] For example, see Ofwat (2024), 'Base cost adjustment claim feeder model – South West Water', found here https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-SWB_Cost-adjustment-claims.xlsx, last accessed 6 August 2024.
[15] It is possible that the regression outputs may not be fully aligned with operational expectations, given that these models do not necessarily capture specific 'service drivers'. This is to be expected, given that Approach 2 is not a fully integrated model. Nonetheless, under Approach 2, the cost drivers will implicitly capture the differences in service across companies based on the statistical correlation in the data and not the operational relevance of certain cost drivers.

amendments to its model specification for the FDs, as this would naturally have an impact on the implicitly funded level of performance (unlike under Approach 1). More importantly, Approach 2 does not specifically account for the drivers of service performance when determining what a reasonable target is—if there are drivers of service performance that are not reflected in the cost assessment models (such as weather and climate), then the implicitly funded level of performance under Approach 2 may set overly stringent or relaxed targets for individual companies.

Approach 3: predicted performance—service drivers

Ofwat's approach to setting PCs does not generally account for the specific drivers of service performance. Therefore, any comparative assessment used to construct the PCs assumes that companies operate in similar operating environments, such that the performance of one company can be replicated by another. This is inconsistent with Ofwat's approach to setting base allowances, where it explicitly controls for regional characteristics (such as density) when making comparisons across companies. That is, Ofwat does not assume that the cost (or unit cost) achieved by one company can simply be replicated by another.

Some companies had proposed accounting for regional factors when setting PCLs in their business plans.[16] However, Ofwat has generally argued that this would amount to double-counting, since the models already fund companies to deliver the same level of service. As noted with Approach 2, this is an incorrect argument—companies will be funded for a different level of service depending on the correlation between service quality and the cost drivers in the models, in the same way that Ofwat argues that (for example) companies are funded for a different level of regional wages depending on the density of their operating environments.

Moreover, Ofwat has applied this logic inconsistently when setting PCLs. For example, Ofwat argues that its cost models already account for scale and density and, as such, companies are funded to deliver the same level of supply interruptions regardless of their scale and density—indeed, Ofwat argues that accounting for scale and density when setting the PCL would amount to a double-count.[17] However, Ofwat *does* account for scale when setting the PCL, given that the PCL is defined by

[16] For example, see Oxera (2023), 'Econometric modelling of sewer flooding performance', September, prepared for Dŵr Cymru.
[17] See Ofwat (2024), 'PR24 draft determinations: Delivering outcomes for customers and the environment', July, p. 102.

the average minutes lost *per property* and not the total minutes lost. Applying Ofwat's logic, it is not clear why accounting for density results in a double-count, but accounting for scale does not.

In the same way that Ofwat develops econometric models to assess companies' base cost requirements (and some enhancement expenditure requirements), it is helpful to develop econometric models to assess companies' PCs. Unlike with Approach 2, there is an explicit recognition that there are likely to be drivers of service performance that are not accounted for in Ofwat's cost assessment modelling, and Ofwat would be required to develop service performance models using a broadly similar model development approach to the one that it has adopted for the base cost modelling.

A limitation with this approach, which is also embedded in Ofwat's current approach, is that there is no integration between cost and service performance. If one company performs well in the cost models and sets the benchmark, yet performs poorly in the service models, and another company performs well in the service models and sets the benchmark, but performs poorly in the cost models, the overall stretch on cost and outcomes may be unachievable. Therefore, care must be taken when setting the benchmark in both the cost and the service models. One ad hoc method of combining the assessment would be to set the benchmark in the service models based on the performance of the cost-efficient companies (or to set the benchmark in the cost models based on the performance of the service-efficient companies).

Approach 4: integrated analysis

The final approach that we have considered is a full integration of the assessment of cost and service quality. We consider that this could be achieved in one of three ways, as follows.

The conceptually simplest method of integrating cost and quality is to include measures of service quality within the cost assessment models. Here, the relationship between cost and output is estimated directly by the models, such that the expected costs for achieving a certain level of service (such as Ofwat's PCs) can be estimated by extrapolating the cost models into the future. This approach has been adopted by several regulators across sectors and jurisdictions.[18] While this is conceptually

[18] For example, ARERA (the Italian water regulator) controls for service measures such as reliability in its cost assessment models (see ARERA (2022), 'Regulation and efficiency costs in the Italian water sector', June).

simple and can provide a useful cross-check to other analyses, the relationship between cost and quality is likely to be complex[19] and can be difficult to estimate robustly in an econometric framework on the available dataset.

We note that Professor Andrew Smith, an academic advisor to Ofwat, recommended in his review statement on the PR24 base cost consultation models that Ofwat explore the approach of incorporating measures of service quality in its cost assessment models. Smith suggests that:[20]

> to the extent that quality can successfully be incorporated into regulatory cost models it can give useful information on the cost of quality improvements and allow benchmarking to take into account a wider range of cost drivers. As a longer term objective I would therefore support continued development and consideration of these kinds of approaches going forward alongside other approaches to incentivising quality.

Alternatively, Ofwat could consider developing monetised measures of service quality and adding these to the modelled cost base. This approach is fairly common in the energy sector, where some regulators monetise supply interruptions and add this to modelled TOTEX when assessing performance.[21] This approach mitigates the complications associated with modelling complex relationships in an econometric framework. The approach to monetising service could be driven by the expected costs associated with improving performance (which could be derived through bottom-up analyses) or the consumer benefit of improving service (which could be derived through customer surveys), or some combination of the two.

Lastly, Ofwat could consider systems modelling. This would involve developing models for costs and service (the service models would be developed in line with Approach 3) and estimating such models jointly using (for example) seemingly unrelated regressions. Such an approach can explicitly account for the common and separate drivers of cost and service, and can also model the trade-off between cost and service

[19] For example: (i) there may be a lagged relationship between expenditure and service, such that expenditure in one year does not necessarily affect service quality in that year but, instead, affects service in future; (ii) the marginal cost of improving service may depend on regional factors (including a company's current level of service); and (iii) there are several different measures of service, such that some form of aggregation may be required.
[20] See Ofwat (2023), 'Econometric base cost models for PR24', April, section 5.
[21] For example, see Oxera (2020), 'Quality measures in cost benchmarking'.

directly. This approach is more sophisticated and academically rigorous than the approaches outlined above.

Examples of implementation

We have explored the extent to which some of Ofwat's PCLs are implicitly funded through the base expenditure models under the approaches described above. As noted, Approaches 1 and 2 can be applied quickly, while Approaches 3 and 4 require a more intensive model development exercise that has not been possible to complete robustly given the tight deadlines imposed for responding to the DDs. Therefore, the figures presented for Approach 3 should be considered as preliminary, and we have currently not developed models under Approach 4.

The table below shows how YWS's PCL for 2030 compares with the implicitly funded level of service according to these approaches.

Implicitly funded PCL for 2029/30

	Ofwat DD	Approach 1	Approach 2	Approach 3
Supply interruptions (hh:mm:ss)	00:05:00	00:15:30	00:11:57	00:14:01
Leakage (Ml/d)	224	268	261	269
Internal sewer flooding (incidents per 10,000km)	1.16	2.19	2.37	2.49
Total pollution incidents (incidents per 10,000km)	13.65	38.32	39.00	27.88

Source: Oxera analysis.

For all of the PCs we have explored, the PCLs set by Ofwat are materially more stringent than what is funded through the models. Given that YWS has not received sufficient[22] additional funding to achieve these stretching PCs, this amounts to a large and unjustified efficiency challenge for YWS. While we have not explored other PCs in detail as

[22] YWS has received some additional funding through Ofwat's post-modelling adjustments that may enable it to improve its performance on some PCs (for example, meter renewals may improve leakage, and mains replacement may improve mains repairs), but these post-modelling adjustments do not explicitly account for the level of service that YWS is expected to achieve. This means that YWS is likely to be underfunded.

part of this review, we consider that it is probable that the PCLs for other PCs are also likely to be overly stretching, given that:

- the year 0 PCL is typically based on the PR19 PCL, which (as noted above) was not explicitly funded at PR19 and companies have generally underperformed on service in AMP7;
- for several PCLs, Ofwat has imposed an additional stretch on performance beyond the year 0 PCL.

To address this issue at the FD, Ofwat should consider more robust approaches to investigate the relationship between cost and service. Following the approaches outlined in this report, Ofwat should consider either relaxing some of the PCLs where the performance is clearly unfunded, or providing additional allowances for companies to achieve the stretching PCLs, or some combination of the two. Otherwise, the overall package offered at PR24 is unlikely to be achievable.

A1 Assessing the PR19 approach

Ofwat did not integrate the assessment of cost and outcomes at PR19. Instead, it set separate targets on the basis of separate models for each area. The costs were determined through a combination of econometric modelling (for base expenditure and some enhancement categories) and bottom-up assessments, while PCLs were in general assessed using one or more of the following approaches:[23]

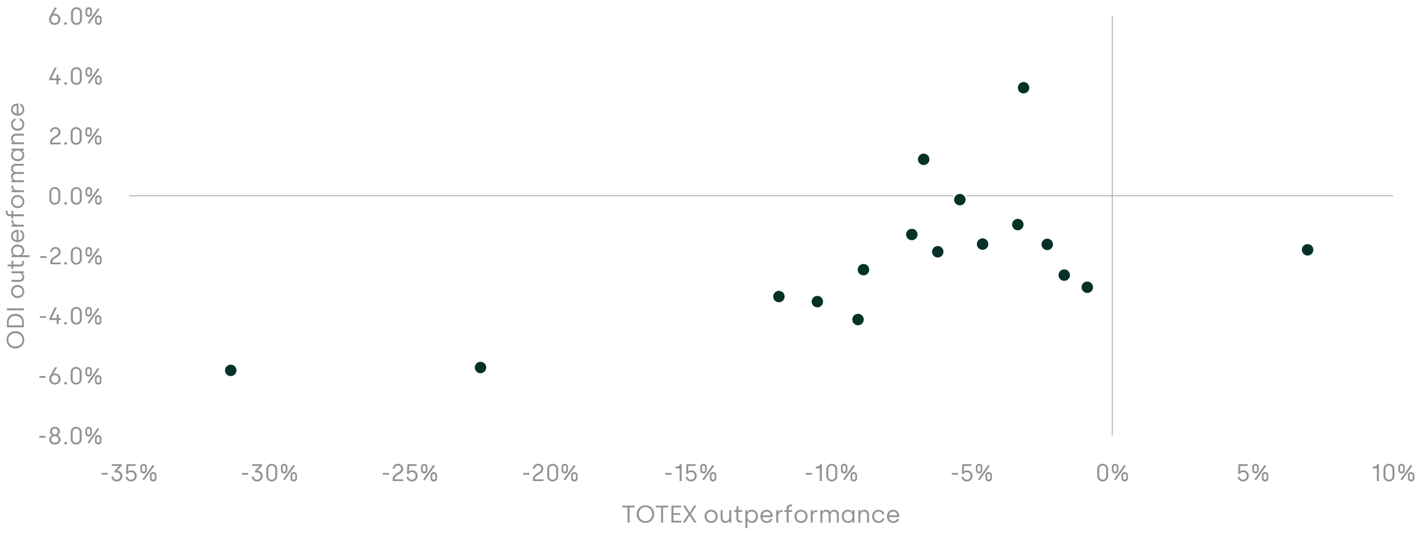
- looking at the UQ performance forecast in companies' business plans (for example, in terms of internal sewer flooding or pollution incidents);
- assuming a specific rate of improvement relative to the PR14 PCL (for example, in terms of leakage);
- ensuring full compliance with statutory measures (for example, water quality compliance);
- looking at the UQ rate of improvement proposed in companies' business plans (for example, in terms of sewer collapses);
- calculating the average of a company's 'best-performing' outturn years (for example, in terms of mains repairs).

In some cases, Ofwat made company-specific adjustments to the PCL to reflect unique operating environments.

While Ofwat did not present robust evidence that the cost and service targets were simultaneously (i.e. 'in the round') achievable at PR19, the targets may have nonetheless been achievable if there were some hidden sources of efficiency gains available to companies. In this case, we would expect companies to have generally met or surpassed these targets during AMP7, given the strong incentives for companies to improve performance while reducing costs. The figure below shows companies' outperformance levels on cost and service to date.

[23] Ofwat (2019), 'PR19 final determinations: Delivering outcomes for customers policy appendix', December.

Figure A1.1 Outperformance throughout AMP7 (2021–23)



Source: Oxera analysis of Ofwat (2024), 'Data for the Water Company Performance Report 2022-23', February.

If the PR19 determination represented a 'fair bet', one would expect the companies to be clustered around the origin, with individual companies equally likely to outperform or underperform on costs and service. Indeed, we might expect most companies to fall either in the top-left quadrant or the bottom-right quadrant, depending on their management focus. If the PR19 determinations were 'challenging but achievable', companies that placed a particular focus on improving service quality could only outperform the service targets by spending more than their allowances and underperforming on costs (top-left quadrant), or vice versa (bottom-right quadrant).

However, we find that the majority of companies are in the bottom-left quadrant, underperforming on both cost and service quality. Indeed, no company is in the top-right quadrant—i.e. no company has been able to exceed both its cost and its quality targets. This supports companies' concerns with the PR19 determination that the overall stetch on cost and performance was unachievable. That is, companies were not sufficiently funded to deliver the stretching PCs or, conversely, the PCs were too stretching to be delivered within the given expenditure allowances.

Given that companies have not been funded to achieve the PR19 PCs, using the PR19 PCs as a baseline target for 'year 0' in PR24 (from which companies are generally expected to improve performance further) represents a material and unsubstantiated efficiency challenge.

In the PR24 methodology, Ofwat stated that it would undertake a more detailed assessment of the link between cost and service quality. However, the evidence presented in the DD appears to follow the PR19 approach in places, and does not adequately address the challenges with the PR19 determination.

A2 Examining companies' proposed targets

At both PR19 and PR24, Ofwat has used companies' proposed service improvements either to directly determine companies' PCLs or as a cross-check to whether the PCLs are achievable. The logic behind this approach is that, if at least some companies are proposing to deliver a stretching PCL, it should be feasible for all companies to deliver that PCL. However, this logic is flawed for several reasons.

First, companies are strongly incentivised to propose 'ambitious' (i.e. low-cost, high-service) plans. Doing so provides companies with reputational and financial rewards, while failing to do so results in reputational and financial penalties. It is feasible that at least some companies will have proposed 'overly ambitious' and ultimately unachievable business plans. Relying narrowly on the information included within some companies' business plans to inform PCLs imposes a risk on the rest of the sector that the PR24 determination will be unachievable.

Ofwat used a forward-looking benchmark (i.e. the proposed stretch in companies' business plans) to inform companies' residential retail allowances at PR19. Some of the companies had proposed material efficiency improvements in their business plans, such that companies that were estimated to be inefficient (sometimes materially) on an outturn basis set the benchmark on a forward-looking basis. However, these efficiency improvements did not materialise in AMP7, and the overall level of challenge proved to be ultimately unachievable, given that companies overspent their allowances by c. 20% on average.[24] This is an example of the risks associated with relying on forward-looking data for a handful of companies to set targets for the rest of the sector.

Second, companies do not propose service targets in isolation but as part of an overall package for consumers. Companies would expect to meet the service targets only if they are also able to spend what they have proposed (on both base and enhancement). However, Ofwat has made material challenges to companies' proposed expenditure, such that it cannot be assumed that companies can deliver their proposed targets within the expenditure allowances. That is, companies are not

[24] Indeed, we note that only one company, Hafren Dyfrdwy (HDD), has materially underspent its allowance (by c. 19%). However, Severn Trent England (SVE), operating under the same ownership group as HDD, overspent its allowance by 11%, such that the combination of SVE and HDD (SVH) underperformed. See Ofwat (2023), 'Water company performance report 2022–23', September, slide 30.

funded to deliver the service levels proposed in their plans, but funded to deliver the service levels that are implicitly funded through Ofwat's models (and adjustments thereof).

Similarly, a company may propose a more stretching service target in one area and a more relaxed target in another (for example, because of customer preferences or long-term objectives). Using one company (or set of companies) to inform the target in one service area and another company (or set of companies) to inform the target in another could amount to cherry-picking, and could result in an unachievable determination in the round. We note that not a single company has had its proposed service levels accepted in full.[25]

Third, Ofwat's comparisons of companies' proposed service targets are overly simple and do not account for relevant regional factors that may drive performance. That is, some companies may be better able to achieve stretching PCLs than others on account of their operating environment (for example, density, topography or climate) or historical enhancement expenditure allowances.

[25] Oxera analysis of Ofwat's DD and companies' business plan submissions.

A3 Rate of improvement or level of service?

When determining what is implicitly funded through base allowances, Ofwat assumes that companies can sustain their rate of improvement in service through base allowances rather than sustain the level of service for some PCLs only (for example, storm overflows). However, this assumption is inconsistent with how companies deliver performance improvements in practice.

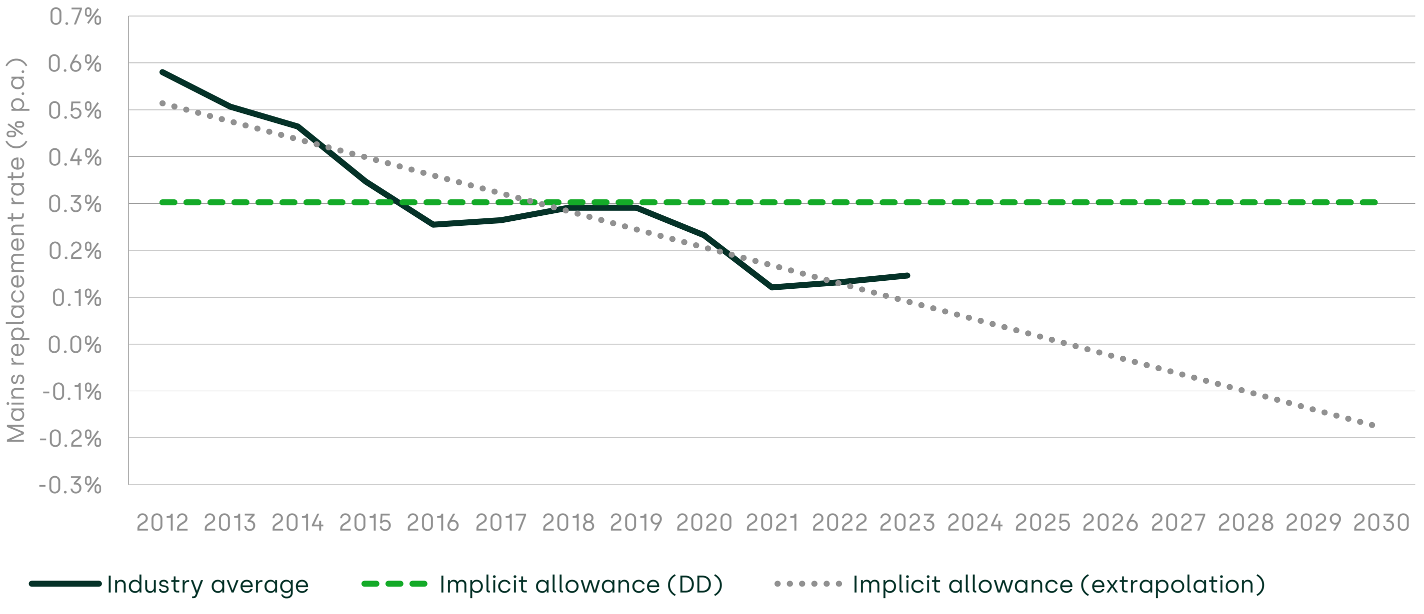
For example, a company may invest in maintenance activities in order to improve asset health and improve associated PCs (such as mains bursts, leakage or storm overflows). Once the expenditure is incurred, the assets are in a better condition and a company's performance will improve to a new level. However, in order to improve beyond that new level, a company would need to further improve the health of its assets or find other solutions to improve performance (which would be associated with additional—albeit possibly different—costs).

Similarly, a company seeking to reduce leakage may invest in the monitoring of its network. As the network is monitored more regularly, failures on the network can be identified and fixed more quickly, resulting in less leakage. However, if a company intended to improve leakage further, it would need to spend even more on monitoring the network.

That is, it cannot be assumed that the outturn trend in service improvements can be simply extrapolated into the future.

In Ofwat's post-modelling adjustments, it does not assume that the trend in activity is implicitly funded in the models. Instead, it assumes that the average level of activity is implicitly funded. We note that, if Ofwat were to apply its logic that the *trend* is funded and not the *level*, this could lead to somewhat perverse outcomes. For example, the figure below shows how the implicitly funded rate of mains replacement activity estimated via a trend compares with the value that Ofwat has determined.

Figure A3.1 Mains replacement extrapolation



Source: Oxera analysis.

The trend in mains replacement activity across the industry would suggest that companies are implicitly funded to deliver negative mains replacement (c. -0.2% p.a.) by 2030, using Ofwat's logic.

A4 Estimating the implicitly funded level

In setting the level of service that should be implicitly funded through the cost models, one has to make assumptions regarding how service affects costs and how service is related to the cost drivers included in Ofwat's cost assessment models. In essence, the issue amounts to an 'omitted variable' problem—that is, there is some driver of expenditure (i.e. service quality) that is related to expenditure 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.

Under Approach 1, consistent with Ofwat's approach, we assume that service quality is uncorrelated with the cost drivers included in the models. In this case, service quality can be treated as a random variable. Suppose that Ofwat's models are otherwise unbiased and that there are no other omitted factors. The true cost function is:

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

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;
- $Service\ quality_{it}$ is the observed service quality 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 service quality 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 service quality over the modelling period—i.e. the implicitly funded level of service. In this stylised case, it would be broadly appropriate to determine the implicitly funded level of service as the industry-average activity over the modelling period (i.e. Ofwat's approach for estimating the implicitly funded level of activity in its post-modelling adjustments).

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 service quality. If, instead, there is a strong correlation between the cost drivers and the omitted factor, 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 service quality, such that the implicitly funded level of activity would differ by company depending on the value of that cost driver (this is discussed in more detail under Approach 2 below). Nonetheless, assuming that service quality is uncorrelated with the cost drivers may be an appropriate and proportionate simplifying assumption in some cases.

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 benchmark 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 unadjusted output from the regression. Instead, while the coefficients of the cost drivers are indeed the unadjusted 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 service 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 (or, potentially, the average activity of the cost-efficient companies). 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 (or companies), 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 by management, given prior flexibility on what companies were able to direct funding to;
- relying specifically on the cost-efficient companies 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.

Under Approach 2, we relax (or rather test) the assumption that service quality is uncorrelated with the cost drivers included in the models. In this case, the cost impact of service is partially accounted for in the estimated coefficients in the models, and therefore companies are funded for a different level of service depending on their operating environment. We note that Ofwat applies this logic when rejecting several companies' CACs, including as follows.

- Regional wages: Ofwat has rejected claims relating to regional wages, in part arguing that the cost impact of regional wages is captured by the density drivers in the cost models.[26] That is, companies are funded for a particular level of regional wages depending on their level of density.
- WTW-level economies of scale: Ofwat has rejected claims relating to WTW-level economies of scale, in part arguing that the cost impact of WTW-level economies of scale is implicitly captured by the density drivers in the cost models.[27] That is, companies are funded for a particular average WTW size depending on their level of density.
- Liming and bioresources: Ofwat has rejected a claim relating to the additional costs associated with particular treatment technologies, arguing that treatment technology is correlated with treatment work size.[28] That is, companies are implicitly funded for particular treatment technologies depending on the size of their treatment works.

While Ofwat uses these arguments to suggest that the cost impact of various omitted factors is already implicitly captured by the models, it does not undertake a thorough assessment of *the extent to which* these omitted factors are implicitly funded. The assessment is binary—either

[26] For example, see Ofwat (2024), 'Base cost adjustment claim feeder model – Affinity Water', found here https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-AFW_Cost-adjustment-claims.xlsx, last accessed 6 August 2024.
[27] For example, see Ofwat (2024), 'Base cost adjustment claim feeder model – South East Water', found here https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-SEW_Cost-adjustment-claims.xlsx, last accessed 6 August 2024.
[28] For example, see Ofwat (2024), 'Base cost adjustment claim feeder model – South West Water', found here https://www.ofwat.gov.uk/wp-content/uploads/2024/07/PR24-DD-SWB_Cost-adjustment-claims.xlsx, last accessed 6 August 2024.

the omitted factor is captured in the models (in which case, Ofwat assumes that the full cost impact of the omitted factor is captured in the models) or the omitted factor is not captured in the models (in which case, Ofwat would presumably assume that an adjustment is required). If we were to apply the same binary logic to service quality, the implicitly funded level of service for a given company would be what that company has historically achieved.

However, we do not consider that this binary reasoning is appropriate in the context of determining what level of service is implicitly funded. At best, the high-level cost drivers included in Ofwat's models would capture the costs associated with service quality imperfectly, such that it cannot be assumed that the presence of any correlation between service quality and the cost drivers means that companies are fully funded for all differences in service levels. Instead, we consider that it is important to quantify the extent to which service quality is correlated with the cost drivers.

In the context of PCs, we consider that the extent to which the cost drivers capture differences in service (and therefore implicitly fund different levels of service) could be estimated through the following procedure.

1. For each relevant cost model, regress the service measure against the cost driver included in that model.
2. Use these models to predict the level of service for each company.
3. Triangulate the predicted level of service across the models, following Ofwat's triangulation approach.

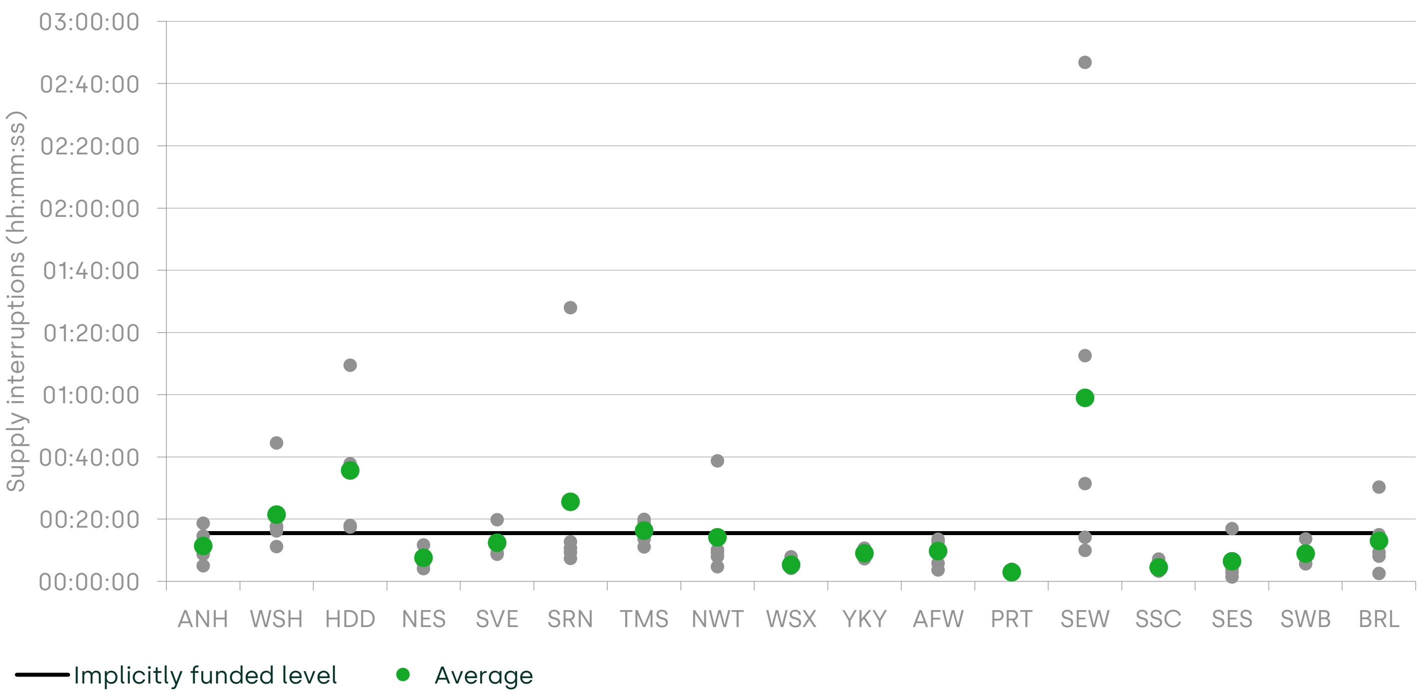
A5 Supply interruptions

Ofwat has set a common target of an average supply interruption of five minutes, to be achieved in each year of AMP8. This is consistent with companies' targets for 2024/25 that were set at PR19, where Ofwat used a combination of historical trend analysis and direct comparisons to set the common target of five minutes. To justify the five-minute target at PR24, Ofwat argues that: (i) nine out of 17 companies are expecting to meet or exceed this target in 2025; and (ii) this target is approximately equal to the median forecast target provided by companies in 2030, with 14 out of 17 companies expecting to meet or exceed this target.

A5.1 Approach 1: industry average

The figure below shows the implicitly funded level for average supply interruptions in the benchmark period (2019–23). Each grey dot represents a company's value in a given year, while the green dot represents the average duration for that company in the benchmark period.

Figure A5.1 Water supply interruptions (2019–23)



Note: Each grey dot represents the performance of a given company in a given year, while the large green dot represents the average performance of that company over the benchmark period (2019–23).

Source: Oxera analysis.

The figure shows that companies are implicitly funded to deliver c. 15 minutes and 30 seconds of supply interruptions, based on this approach. This is materially less stringent than Ofwat's target of five minutes.

Note that the historical average may be skewed by the performance of some companies. While we do not consider that it would be appropriate to adjust the methodology in light of this observation,[29] we note that the median duration of supply interruptions over the benchmark period is c. 11 minutes and 20 seconds, which is still materially less stringent than Ofwat's target.

A5.2 Approach 2: predicted service—cost drivers

The table below shows the correlation between supply interruptions (in terms of seconds lost, in logs) and the cost drivers included in Ofwat's treated water distribution (TWD) and wholesale water (WW) models.

Table A5.1 Water supply interruptions—correlation table

	Seconds lost (log)
Length of mains (log)	0.229***
Boosters per mains (log)	0.353***
Weighted average density, MSOA to LAD (log)	-0.300***
Weighted average density, MSOA to LAD (log), squared	-0.291***
Weighted average density, MSOA (log)	-0.249***
Weighted average density, MSOA (log), squared	-0.237***
Properties per length of mains (log)	-0.294***
Properties per length of mains (log), squared	-0.289***
APH TWD (log)	0.270***
Connected properties (log)	0.144*
Proportion of water treated in complexity bands W3–6	0.228***
Weighted average complexity (log)	0.266***

[29] Strictly speaking, companies would be implicitly funded based on the mean of the observations included in the sample under the observed assumptions, such that the presence of outliers influences the cost models and the implicitly funded level of performance.

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

As shown in the table, supply interruptions is correlated with some of the cost drivers included in Ofwat's models, particularly some measures of density and topography. While the correlations are statistically significant, the relationship correlation is not particularly strong (the correlation coefficient is less than 0.3). However, these correlations are partial and do not take into account the fact that the combination of cost drivers may be better correlated with supply interruptions.

The tables below show the estimated relationship between supply interruptions and the cost drivers included in Ofwat's DD models.

Table A5.2 Supply interruptions: TWD models

	TWD1	TWD2	TWD3	TWD4	TWD5	TWD6
Length of mains (log)	0.242***	0.232***	0.243***	0.228***	0.220***	0.234***
Boosters per mains (log)	0.748***	0.691***	0.731***			
Weighted average density, MSOA to LAD (log)	-2.086*			-3.266***		
Weighted average density, MSOA to LAD (log), squared	0.142*			0.215***		
Weighted average density, MSOA (log)		-9.299***			-12.61***	
Weighted average density, MSOA (log), squared		0.572***			0.772***	
Properties per length of mains (log)			-12.11*			-19.32***
Properties per length of mains (log), squared			1.387*			2.188***
APH TWD (log)				0.640***	0.611***	0.647***
Constant	14.93***	44.84***	33.60**	13.76***	53.07***	43.93***
Observations	164	164	164	164	164	164
R-squared	0.208	0.233	0.210	0.218	0.233	0.219

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependant variable is the natural logarithm of seconds lost.

Source: Oxera analysis.

Table A5.3 Supply interruptions: WW models, booster pumping stations

	WW1	WW2	WW3	WW4	WW5	WW6
Connected properties (log)	0.180***	0.156**	0.187***	0.169**	0.191***	0.166**
Boosters per mains (log)	0.706***	0.721***	0.767***	0.758***	0.789***	0.773***
Proportion of water treated in complexity bands W3–6	0.0108***		0.00848***		0.00894***	
Weighted average complexity (log)		1.093***		0.904***		0.989***
Weighted average density, MSOA to LAD (log)	-2.775**	-2.092*				
Weighted average density, MSOA to LAD (log), squared	0.184**	0.136*				
Weighted average density, MSOA (log)			-9.334***	-8.284***		
Weighted average density, MSOA (log), squared			0.570***	0.503***		
Properties per length of mains (log)					-12.72**	-10.51*
Properties per length of mains (log), squared					1.441**	1.179
Constant	16.43***	13.63***	44.51***	39.90***	34.39***	29.25**
Observations	164	164	164	164	164	164
R-squared	0.262	0.259	0.270	0.270	0.249	0.253

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependant variable is the natural logarithm of seconds lost.
Source: Oxera analysis.

Table A5.4 Supply interruptions: WW models, APH

	WW7	WW8	WW9	WW10	WW11	WW12
Connected properties (log)	0.178***	0.163**	0.189***	0.181**	0.204***	0.190***

	WW7	WW8	WW9	WW10	WW11	WW12
APH TWD (log)	0.434**	0.431**	0.474**	0.462**	0.533**	0.503**
Proportion of water treated in complexity bands W3–6	0.00822**		0.00446		0.00458	
Weighted average complexity (log)		0.777**		0.454		0.527
Weighted average density, MSOA to LAD (log)	-3.645***	-3.137***				
Weighted average density, MSOA to LAD (log), squared	0.235***	0.200***				
Weighted average density, MSOA (log)			-12.87***	-12.32***		
Weighted average density, MSOA (log), squared			0.779***	0.744***		
Properties per length of mains (log)					-19.51***	-18.18***
Properties per length of mains (log), squared					2.182***	2.025***
Constant	15.35***	13.28***	54.35***	52.02***	44.37***	41.49***
Observations	164	164	164	164	164	164
R-squared	0.243	0.237	0.238	0.236	0.228	0.229

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependant variable is the natural logarithm of seconds lost.
Source: Oxera analysis.

The model fit in these regressions is fairly low (always less than 30%), which may indicate that there is not a strong relationship between supply interruptions and the cost drivers included in Ofwat's cost models. However, nearly all of the cost drivers are statistically significant, indicating that the estimated coefficients in Ofwat's cost models may implicitly capture some of the costs associated with achieving different levels of supply interruptions. That is, companies will be implicitly funded for different levels of supply interruptions, depending on their operating environment.

The table below shows how YWS's implicitly funded level of performance via this approach compares with Ofwat's PC.

Table A5.5 Supply interruptions: Approach 2 target

	2024/25	2025/26	2026/27	2027/28	2028/29	2029/30
Ofwat proposed	00:05:00	00:05:00	00:05:00	00:05:00	00:05:00	00:05:00
Approach 2	00:12:01	00:12:00	00:11:59	00:11:59	00:11:58	00:11:57

Source: Oxera analysis.

While the implicitly funded level of supply interruptions is lower under Approach 2 compared to Approach 1 (c. 15 minutes and 30 seconds), it is significantly less stringent than Ofwat's target, indicating that YWS is underfunded to deliver on this improvement.

A5.3 Approach 3: predicted service—service drivers

At this stage, we consider that the following factors may drive supply interruptions.

- 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.
- Topography—companies that require more network assets due to topography may have more (and longer) supply interruptions due to the having more 'points of failure'.
- Asset health—assets that are in better condition are less likely to fault and therefore less likely to result in a supply interruption.
- Water scarcity and resilience—companies that have high levels of resilience (for example, due to water abundance, historical asset configuration, and historical and current enhancement allowances) may be better able to resolve faults on the network without causing supply interruptions.

Given the lack of data on relevant drivers (such as water scarcity and resilience), we have estimated simple models that control for density, topography and asset health, as shown in the table below.

Table A5.6 Supply interruptions: performance modelling

	Supply 1	Supply 2	Supply 3	Supply 4	Supply 5	Supply 6
Percentage of metered households		-0.0173**		-0.0159*		-0.0150*

	Supply 1	Supply 2	Supply 3	Supply 4	Supply 5	Supply 6
Percentage of mains in condition grades 4 and 5	6.184***	6.557**	7.067***	7.372***	8.864***	9.295***
Weighted average density MSOA (log)					-1.000***	-1.161***
Properties per length (log)			-1.573***	-1.903***		
Weighted average density LAD (log)	-0.455***	-0.594***				
Constant	9.607***	11.55***	13.04***	15.34***	14.15***	16.24***
RESET	0.00841	2.52e-09	0.0126	0.269	0.00198	0.490
R-squared	0.161	0.0814	0.171	0.0992	0.170	0.0970

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the natural logarithm of the average seconds lost.
Source: Oxera analysis.

We note that all drivers are statistically significant at the standard thresholds, and have a directionally intuitive coefficient. However, the model fit is low when compared with Ofwat's equivalent models for expenditure. While the models fail the RESET test, we have explored included squared density terms but the estimated coefficients are statistically insignificant and/or unintuitive.

While there are limitations with the models presented above, these may provide a reasonable starting point for assessing the achievability of Ofwat's PCs. We reiterate that this is a preliminary selection of the drivers of supply interruptions, and Ofwat should consult with the industry when developing performance models as per its current approach to developing the cost models. It may be appropriate to combine top-down econometric modelling with more bottom-up, engineering assessments to determine what is feasible within base expenditure.

The table below shows the implicitly funded level of supply interruptions under this approach compared with Ofwat's PC.

Table A5.7 Supply interruptions: Approach 3

	Supply 1	Supply 2	Supply 3	Supply 4	Supply 5	Supply 6	Triangulated
2026	00:16:49	00:15:36	00:15:35	00:14:05	00:16:51	00:15:26	00:15:44
2027	00:16:34	00:15:01	00:15:17	00:13:32	00:16:28	00:14:47	00:15:16
2028	00:16:18	00:14:28	00:14:59	00:13:00	00:16:06	00:14:11	00:14:50
2029	00:16:03	00:13:57	00:14:42	00:12:30	00:15:44	00:13:36	00:14:25
2030	00:15:48	00:13:26	00:14:26	00:12:02	00:15:23	00:13:03	00:14:01

Source: Oxera analysis.

The triangulated predicted performance for YWS is c. 15 minutes and 44 seconds in 2026, and this declines to c. 14 minutes by 2030. The predicted improvement in service is driven largely by an improvement in YWS's asset health, which is a function of Ofwat's post-modelling adjustment for mains replacement activity. That is, even factoring in the additional funding that Ofwat has provided to YWS, the predicted level of performance is materially less stringent than Ofwat's PCL (five minutes).

A6 Leakage

Ofwat's leakage targets are based on companies achieving the PR19 PCL in Year 0 and achieving the leakage targets that companies proposed in their business plans by 2030. Ofwat argues that this is aligned with long-term leakage targets and the reductions implied through the water resources management planning (WRMP) process.

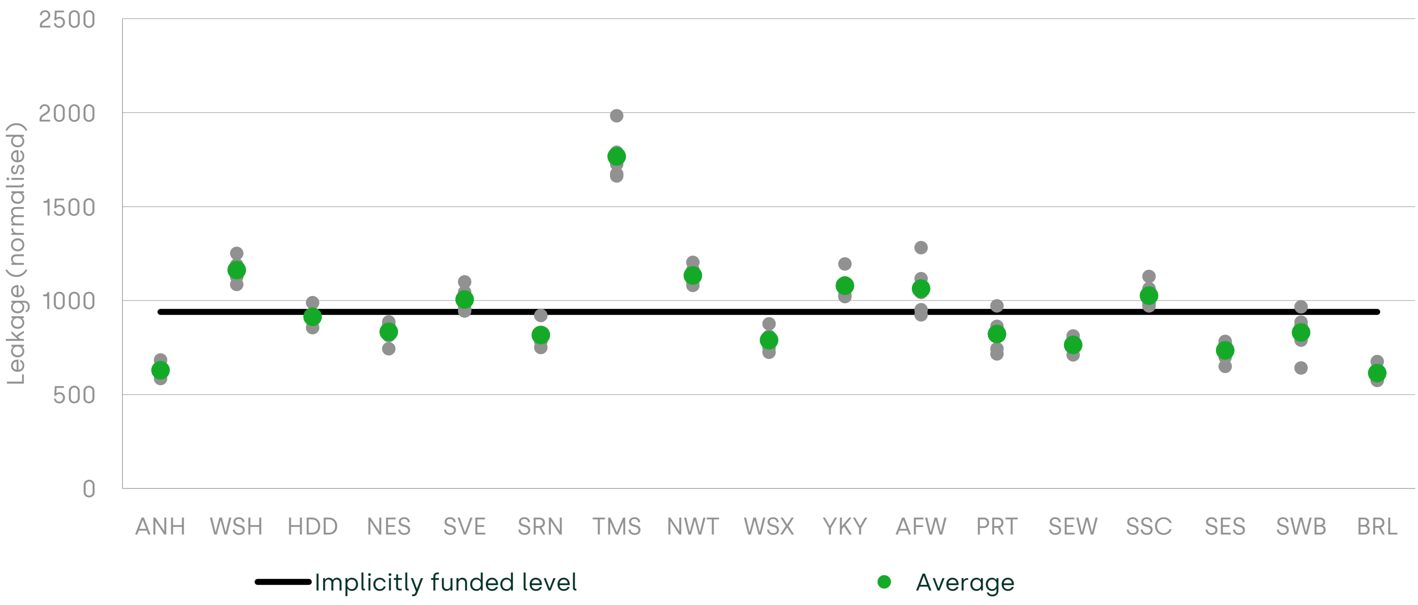
We note that Ofwat provided some companies with additional allowances to fund stretching leakage targets.

A6.1 Approach 1: industry average

In assessing what is implicitly funded through the models under Approach 1, we normalise the level of leakage that companies have achieved (in Ml/d) by the geometric mean of connected properties and length of mains. This is consistent with Ofwat's approach to assessing leakage at PR19.

The figure below shows how companies have performed on this measure in the benchmark period (2019–23), as well as the implicitly funded level under this approach.

Figure A6.1 Leakage performance (2019–23)



Note: Each grey dot represents the performance of a given company in a given year, while the large green dot represents the average performance of that company over the benchmark period (2019–23).

Source: Oxera analysis.

Under this approach, the implicitly funded level of leakage (in Ml/d) depends on a company's scale (as measured by the geometric mean of connected properties and network length). For YWS, the implicitly funded level of leakage under this approach is c. 264Ml/d in 2026 and c. 268Ml/d in 2030. This is considerably less stringent than Ofwat's PCL of 255Ml/d in 2026 and 224Ml/d in 2030.

A6.2 Approach 2: predicted service—cost drivers

The table below shows the correlation between normalised leakage (in logs) and the cost drivers included in Ofwat's TWD and WW models.

Table A6.1 Leakage—correlation table

	Normalised leakage (log)
Length of mains (log)	0.361***
Boosters per mains (log)	-0.101
Weighted average density, MSOA to LAD (log)	0.407***
Weighted average density, MSOA to LAD (log), squared	0.437***
Weighted average density, MSOA (log)	0.417***
Weighted average density, MSOA (log), squared	0.438***
Properties per length of mains (log)	0.374***
Properties per length of mains (log), squared	0.395***
APH TWD (log)	-0.080
Connected properties (log)	0.443***
Proportion of water treated in complexity bands W3–6	0.121
Weighted average complexity (log)	0.166*

Note: Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Normalised leakage is derived by dividing leakage by the geometric mean of connected properties and network length.
Source: Oxera analysis.

Leakage is strongly correlated with several of the cost drivers included in Ofwat's models, including measures of scale, density and treatment complexity. These correlations are generally statistically significant, and the magnitude of the correlation is fairly high.

The tables below show the estimated relationship between leakage and the cost drivers included in Ofwat's DD models.

Table A6.2 Leakage: TWD models

	TWD1	TWD2	TWD3	TWD4	TWD5	TWD6
Length of mains (log)	0.142***	0.114***	0.147***	0.134***	0.112***	0.144***
Boosters per mains (log)	0.293***	0.146*	0.193**			
Weighted average density, MSOA to LAD (log)	-1.860***			-2.231***		
Weighted average density, MSOA to LAD (log), squared	0.146***			0.167***		
Weighted average density, MSOA (log)		-5.147***			-5.799***	
Weighted average density, MSOA (log), squared		0.341***			0.379***	
Properties per length of mains (log)			-13.09***			-14.61***
Properties per length of mains (log), squared			1.600***			1.761***
APH TWD (log)				0.0741	0.0982	0.0938
Constant	-1.414	11.84***	19.00***	-1.303	13.60***	21.35***
Observations	100	100	100	100	100	100
R-squared	0.598	0.522	0.630	0.548	0.516	0.616

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependant variable is the natural logarithm of normalised leakage (leakage divided by the geometric mean of connected properties and network length).
Source: Oxera analysis.

Table A6.3 Leakage: WW models, booster pumping stations

	WW1	WW2	WW3	WW4	WW5	WW6
Connected properties (log)	0.137***	0.142***	0.120***	0.126***	0.145***	0.152***
Boosters per mains (log)	0.289***	0.309***	0.160*	0.155*	0.195**	0.192**

	WW1	WW2	WW3	WW4	WW5	WW6
Proportion of water treated in complexity bands W3–6	0.00113		0.000621		0.000980	
Weighted average complexity (log)		-0.0989		-0.134		-0.121
Weighted average density, MSOA to LAD (log)	-1.917***	-1.852***				
Weighted average density, MSOA to LAD (log), squared	0.147***	0.143***				
Weighted average density, MSOA (log)			-5.324***	-5.522***		
Weighted average density, MSOA (log), squared			0.348***	0.360***		
Properties per length of mains (log)					-13.19***	-13.44***
Properties per length of mains (log), squared					1.595***	1.625***
Constant	-1.703	-1.699	12.22***	13.17***	18.85***	19.55***
Observations	100	100	100	100	100	100
R-squared	0.594	0.593	0.535	0.538	0.633	0.633

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependant variable is the natural logarithm of normalised leakage (leakage divided by the geometric mean of connected properties and network length).
Source: Oxera analysis.

Table A6.4 Leakage: WW models, APH

	WW7	WW8	WW9	WW10	WW11	WW12
Connected properties (log)	0.127***	0.134***	0.117***	0.128***	0.143***	0.155***
APH TWD (log)	0.0501	0.0993	0.0924	0.151**	0.0893	0.145**
Proportion of water treated in complexity bands W3–6	0.00148		-0.000123		0.000279	
Weighted average complexity (log)		-0.120		-0.300*		-0.269*
Weighted average density, MSOA to LAD (log)	-2.287***	-2.242***				

	WW7	WW8	WW9	WW10	WW11	WW12
Weighted average density, MSOA to LAD (log), squared	0.168***	0.165***				
Weighted average density, MSOA (log)			-6.059***	-6.382***		
Weighted average density, MSOA (log), squared			0.391***	0.412***		
Properties per length of mains (log)					-14.73***	-15.31***
Properties per length of mains (log), squared					1.759***	1.829***
Constant	-1.453	-1.633	14.38***	15.64***	21.32***	22.52***
Observations	100	100	100	100	100	100
R-squared	0.543	0.541	0.524	0.540	0.616	0.628

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependant variable is the natural logarithm of normalised leakage (leakage divided by the geometric mean of connected properties and network length).
Source: Oxera analysis.

The model fit in these regressions is high (c. 52–63%), which indicates that there is a strong relationship between leakage and the cost drivers included in Ofwat's models. That is, companies will be implicitly funded for different levels of leakage, depending on their operating environment.

The table below shows how YWS's implicitly funded level of performance via this approach compares with Ofwat's PC.

Table A6.5 Leakage: Approach 2

	2024/25	2025/26	2026/27	2027/28	2028/29	2029/30
Ofwat DD	255	252	243	236	229	224
Approach 2	255	256	258	259	260	261

Source: Oxera analysis.

A6.3 Approach 3: predicted service—service drivers

Leakage arises when there is some fault on the network that results in a loss of water. Ofwat has stated that leakage is affected by:[30]

- operational strategies (for example, pressure management);
- network characteristics (for example, length of mains);
- asset condition (for example, asset age);
- customer base composition (for example, rural or urban).

Given this overview of the drivers of leakage, we consider that it would be appropriate to control for the following factors when building a leakage model.

- 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.
- Meter penetration—leaks are more likely to be identified and reported by consumers if they are metered, given that their water consumption is monitored.
- Historical expenditure allowances—companies that have received additional funds to reduce leakage (either base or enhancement) may be expected to have improved leakage performance relative to other companies, all else being equal.

The table below presents a sample of models that perform reasonably well against Ofwat's modelling criteria.

Table A6.6 Leakage: service modelling

	Leakage 1	Leakage 2	Leakage 3	Leakage 4
Percentage of metered households	-0.0127***	-0.0132***	-0.0132***	-0.0133***
Percentage of mains in condition grades 4 and 5	3.637***	4.395***	4.139***	3.942***

[30] See Ofwat, 'Supply Standards: Leakage', found here https://www.ofwat.gov.uk/households/supply-and-standards/leakage/, last accessed 12 August 2024.

	Leakage 1	Leakage 2	Leakage 3	Leakage 4
Weighted average density MSOA (log)		-0.130		
Properties per length (log)			-0.197	
Weighted average density LAD (log)				-0.0458
Constant	-6.383***	-5.349***	-5.526***	-6.025***
RESET	0.537	0.706	0.729	0.701
R-squared	0.610	0.618	0.612	0.605

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the natural logarithm of normalised leakage, where normalised leakage is defined as leakage (Ml/d) divided by the geometric mean of connected properties and network length.
Source: Oxera analysis.

The estimated coefficients on metered households and asset health are directionally aligned with operational expectations and statistically significant at the 1% level. The estimated coefficients on population density are not statistically significant at the standard thresholds, although the p-values are fairly close to the 10% threshold (c. 0.11–0.20). While the model fit is lower than in Ofwat's base cost models in wholesale water, note that the dependent variable is already normalised for scale and we expect that scale would capture a significant proportion of the leakage across companies. In this sense, these models are equivalent to Ofwat's residential retail and bioresources models (which are unit cost models). The model fit in these leakage models is materially larger than in Ofwat's bioresources models (c. 0.145–0.256) and is broadly comparable to Ofwat's residential retail models (c. 0.143–0.711).

The table below shows YWS's predicted performance in each year of AMP8.

Table A6.7 Leakage: Approach 3

	Leakage 1	Leakage 2	Leakage 3	Leakage 4	Triangulated
2026	288	297	293	294	293
2027	283	290	287	288	287

	Leakage 1	Leakage 2	Leakage 3	Leakage 4	Triangulated
2028	277	283	281	281	281
2029	271	277	274	275	275
2030	266	271	268	270	269

Source: Oxera analysis.

On a triangulated basis, YWS's predicted leakage performance is expected to reduce from 293Ml/d in 2026 to 269Ml/d in 2030. This is driven largely by the improvement in asset health that Ofwat has assumed YWS can deliver as a result of Ofwat's post-modelling adjustment for mains replacement.

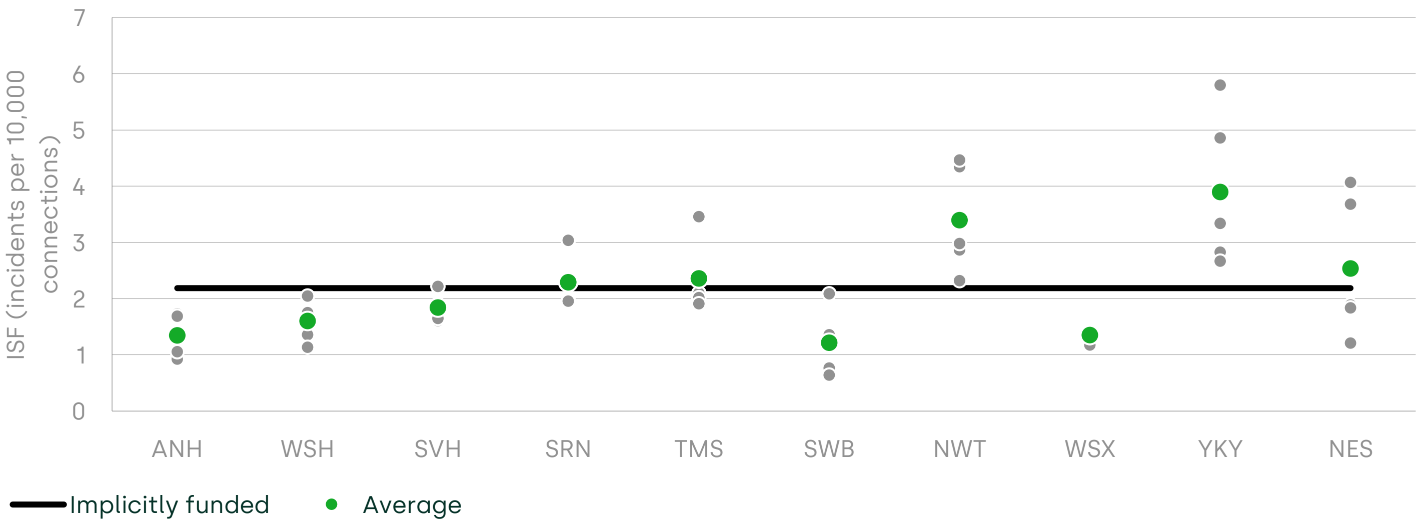
A7 Internal sewer flooding

Ofwat has set a common PCL for internal sewer flooding (ISF) for all companies except Hafren Dyfrdwy (HDD). The year 0 PCL target is set at the PR19 PCL (1.34 incidents per 10,000 sewer connections), which Ofwat argues is achievable given that six companies are forecast to deliver or outperform this level by 2025. Ofwat sets the PCL for 2030 at 1.16 incidents per 10,000 connections, arguing that: (i) this is the median forecast across the industry; and (ii) historical trends suggest that companies can deliver more than this improvement.

A7.1 Approach 1: industry average

The figure below shows the implicitly funded level for ISF, based on the average performance of companies in the benchmark period (2019–23).

Table A7.1 Internal sewer flooding performance (2019–23)



Note: Each grey dot represents the performance of a given company in a given year, while the large green dot represents the average performance of that company over the benchmark period (2019–23).
Source: Oxera analysis.

The figure shows that companies are implicitly funded to deliver c. 2.19 ISF incidents per 10,000 connections, which is significantly higher than both the year 0 PCL target (1.34) and the 2030 PCL (1.16).

A7.2 Approach 2: predicted service—cost drivers

The table below shows the correlation between ISF per 10,000 connections (in logs) and the cost drivers included in Ofwat's sewage collection (SWC) and network plus (WWNP) models.

Table A7.2 Internal sewer flooding—correlation table

	ISF per 10,000 connections (log)
Sewer length (log)	0.243*
Pumping capacity per length (log)	-0.074
Urban rainfall (log)	0.174
Properties per length (log)	0.481***
Weighted average density LAD (log)	0.278**
Weighted average density MSOA (log)	0.354***
Load (log)	0.322**
Percentage load treated at ammonia consents below 3mg/l	0.347***
Percentage load treated at size bands 1–3	-0.573***
Weighted average treatment plant size (log)	0.353***

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

ISF is strongly correlated with most of the cost drivers included in the models. The tables below show the estimated relationship between ISF and the cost drivers included in Ofwat's DD models.

Table A7.3 Internal sewer flooding: SWC models

	SWC1	SWC2	SWC3
Sewer length (log)	-0.179	0.0947	0.00891
Pumping capacity per length (log)	-0.486*	-0.0551	-0.137
Urban rainfall (log)	0.302*	0.428**	0.451**
Properties per length (log)	2.053***		
Weighted average density LAD (log)		0.272**	

	SWC1	SWC2	SWC3
Weighted average density MSOA (log)			0.555***
Constant	-3.908**	-0.962	-2.355
Observations	60	60	60
R-squared	0.313	0.173	0.224

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependant variable is the natural logarithm of normalised internal sewer flooding incidents per 10,000km.
Source: Oxera analysis.

Table A7.4 Internal sewer flooding: WWNP models

	WWNP1	WWNP2
Load (log)	-0.488**	-0.0876
Pumping capacity per length (log)	-0.0690	0.163
Urban rainfall (log)	0.386**	0.332*
Percentage load treated at ammonia consents below 3mg/l	0.00863*	0.00685
Percentage load treated at size bands 1–3	-0.136***	
Weighted average treatment plant size (log)		0.136
Constant	8.383***	1.249
Observations	60	60
R-squared	0.443	0.214

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependant variable is the natural logarithm of normalised internal sewer flooding incidents per 10,000km.
Source: Oxera analysis.

The coefficients on urban rainfall and population density are consistently statistically significant, which may be aligned with operational expectations (see section A7.3 below). The coefficients on load, treatment complexity and economies of scale at the sewage treatment works (STW) level are also statistically significant in some models, which may be unexpected from an operational perspective. Nonetheless, given that there is a strong relationship between ISF and the cost drivers included in Ofwat's models, it is likely that the models fund companies for a particular level of ISF depending on their operating environment.

The table below shows how the implicitly funded level of ISF compares with Ofwat's PC under this approach.

Table A7.5 Internal sewer flooding: Approach 2

	2024/25	2025/26	2026/27	2027/28	2028/29	2029/30
Ofwat DD	1.34	1.31	1.29	1.24	1.2	1.16
Approach 2	2.35	2.35	2.35	2.36	2.36	2.37

Source: Oxera analysis.

The table shows that YWS is implicitly funded to deliver significantly less stretching PCLs under this approach. Assuming that YWS can deliver 1.16 incidents per 10,000km in 2030 without additional funding when it is implicitly funded to deliver only 2.37 incidents per 10,000km results in a material efficiency challenge for YWS.

A7.3 Approach 3: predicted service—service drivers

It is possible to directly model the incidence of internal sewer floodings 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 assets.

As a consequence, we use the following independent variables in our econometric models: 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 A7.6 shows the results from regressing internal sewer flooding incidence, measured as incidents per 10,000 connections, on the percentage of combined sewers and control variables relating to density and urban rainfall.

Table A7.6 Internal sewer flooding: service drivers

	ISF1	ISF2
Combined sewer (percentage)	0.0179*	0.0207**
Urban 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.

We use the results from regressions presented above to predict YWS's incidence of internal sewer floodings over AMP8. Table A7.7 shows Ofwat's proposed PC compared with the results by model as well as the triangulated results.

Table A7.7 Internal sewer flooding: Approach 3

	2024/25	2025/26	2026/27	2027/28	2028/29	2029/30
Ofwat DD	1.34	1.31	1.29	1.24	1.2	1.16
ISF1	2.34	2.33	2.33	2.33	2.33	2.32

	2024/25	2025/26	2026/27	2027/28	2028/29	2029/30
ISF2	2.65	2.65	2.65	2.65	2.65	2.65
Triangulated	2.49	2.49	2.49	2.49	2.49	2.49

Note: Triangulated reflects the average of the values from ISF1 and ISF2.
Source: Oxera.

The table shows that YWS's modelled incidence of internal sewer floodings is materially higher than Ofwat's proposed PC.

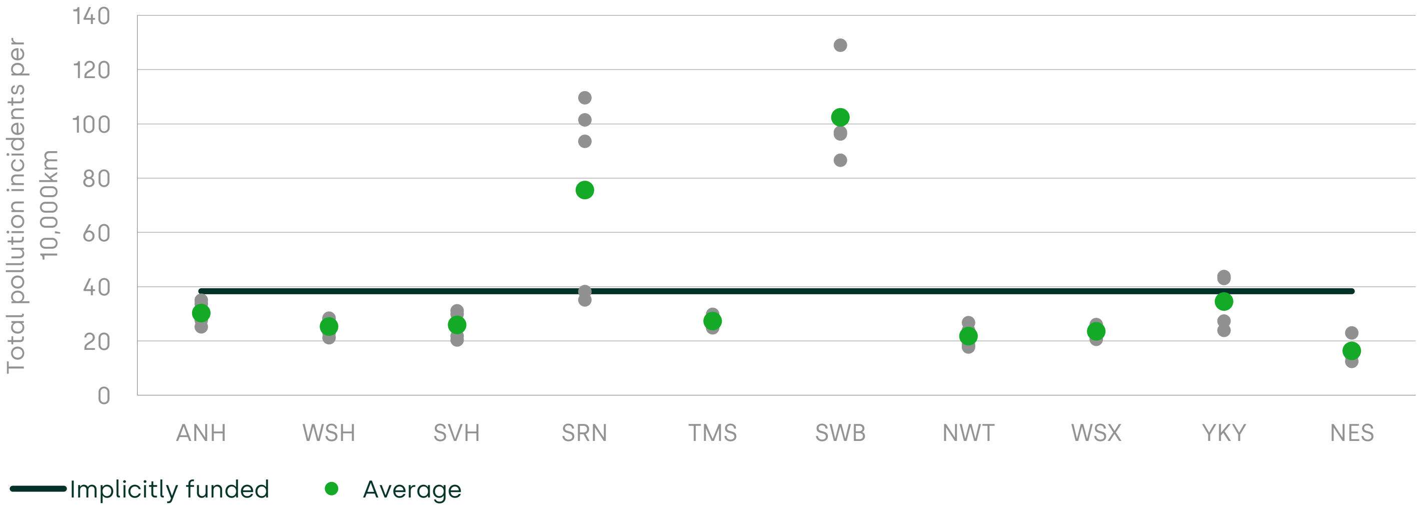
A8 Total pollution incidents

Ofwat has set a common PCL for total pollution incidents for all companies except HDD. The year 0 PCL target is set at the PR19 PCL (19.5 incidents per 10,000km), which Ofwat argues has already been funded. Ofwat sets the PCL for 2030 at 13.65 incidents per 10,000 connections, which represents a c. 30% reduction from the year 0 target over AMP8. Ofwat states that this is aligned with the target set out by the Environment Agency, and that two companies that are currently delivering the PR19 PCL have forecast a 30% improvement.

A8.1 Approach 1: industry average

The figure below shows companies' performance on total pollution incidents during the benchmark period, along with their average performance over the period and the industry average performance (i.e. the implicitly funded level under this approach).

Table A8.1 Total pollution incidents performance (2019–23)



Note: Each grey dot represents the performance of a given company in a given year, while the large green dot represents the average performance of that company over the benchmark period (2019–23).
Source: Oxera analysis.

The figure shows that the implicitly funded level of performance is c. 38 incidents per 10,000km, which is around double the year 0 PCL target and nearly three times the 2030 PCL. Note that the average performance over the benchmark period is somewhat skewed by the performance of two companies, although, as outlined in section A6.1, this should not technically affect the implicit allowance calculations. Nonetheless, even the median performance over the period (c. 27 incidents per 10,000km) is materially less stringent than Ofwat's PCL.

A8.2 Approach 2: predicted service—cost drivers

The table below shows the correlation between total pollution incidents per 10,000km (in logs) and the cost drivers included in Ofwat's sewage treatment (SWT) and WWNP models.

Table A8.2 Total pollution incidents—correlation table

	Total pollution incidents per 10,000km (log)
Pumping capacity per length (log)	0.564***
Urban rainfall (log)	0.090
Load (log)	-0.347**
Percentage load treated at ammonia consents below 3mg/l	-0.208
Percentage load treated at size bands 1–3	0.550***
Weighted average treatment plant size (log)	-0.619***

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

With the exception of urban rainfall and the percentage of load treated at ammonia consents below 3mg/l, there is a strong correlation between the cost drivers and total pollution incidents. The tables below show the estimated relationship between total pollution incidents per 10,000km (in logs) and the cost drivers.

Table A8.3 Total pollution incidents: SWT models

	SWT1	**SWT2**
Load (log)	-0.334	-0.602***
Percentage load treated at ammonia consents below 3mg/l	0.0119**	0.0235***
Percentage load treated at size bands 1–3	0.139***	
Weighted average treatment plant size (log)		-0.644***

	SWT1	SWT2
Constant	6.844*	16.68***
Observations	50	50
R-squared	0.368	0.597

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the natural logarithm of total pollution incidents per 10,000km.
Source: Oxera analysis.

Table A8.4 Total pollution incidents: WWNP models

	WWNP1	WWNP2
Load (log)	-0.322*	-0.706***
Pumping capacity per length (log)	1.346***	0.728***
Urban rainfall (log)	-0.0321	0.119
Percentage load treated at ammonia consents below 3mg/l	0.0195***	0.0267***
Percentage load treated at size bands 1–3	0.163***	
Weighted average treatment plant size (log)		-0.523***
Constant	5.697**	16.78***
Observations	50	50
R-squared	0.734	0.709

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the natural logarithm of total pollution incidents per 10,000km.
Source: Oxera analysis.

The tables show that pollution incidents is strongly associated with nearly all of the cost drivers in Ofwat's DD models, particularly those associated with treatment complexity and STW-level economies of scale. The data also suggests that there is a strong relationship between pumping capacity and total pollution incidents, although it is unclear whether this is aligned with operational expectations.

The table below shows the implicitly funded total pollution incidents per 10,000km under this approach.

Table A8.5 Total pollution incidents: Approach 2

	2024/25	2025/26	2026/27	2027/28	2028/29	2029/30
Ofwat DD	19.5	18.3	17.2	16.0	14.8	13.7
Approach 2	39.7	39.6	39.5	39.2	39.1	39.0

Source: Oxera analysis.

A8.3 Approach 3: predicted service—service drivers

We consider that the following factors may drive pollution incidents.

- Size of sewage treatment works (STW). Since larger STWs handle a larger volume of load, it is possible that these STWs have greater excess capacity to deal with overloads. As such, smaller STWs may be at higher risk of overloads and therefore result in more pollution incidents.
- Scale. Greater load increases pressure on the sewers, which may increase the likelihood of bursts or overflows of sewers.
- Density. As well as the fact that greater volumes of sewage being generated within a given area are likely to increase overloads (as discussed under the size of STWs), discharge is more likely to be categorised as serious in densely populated areas. Discharge in a more dense region is likely to affect a greater number of people and properties, raising the severity of the impact. Rural areas may also have a higher likelihood of a serious pollution incident due to the difficulty of detecting discharge and correcting it, increasing the persistence (time) of the incident.

Given the lack of data on relevant drivers (such as asset health), we have estimated simple models that control for load, weighted average treatment plant size, percentage of load treated at size bands 1–3, and density (which is calculated as the number of properties divided by the total sewer length), as shown in the tables below.

Table A8.6 Total pollution incidents: service drivers

	TPI1	TPI2
Weighted average treatment plant size (log)	-0.646***	
Percentage load treated at size bands 1–3		0.254***

	TPI1	TPI2
Population density (log)	1.955***	3.006**
Constant	2.370	-8.764
Observations	50	50
R-squared	0.538	0.546

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the natural logarithm of total pollution incidents per 10,000km.
Source: Oxera analysis.

The table below shows the implicitly funded total pollution incidents per 10,000km under our augmented approach.

Table A8.7 Total pollution incidents: Approach 3

	2024/25	2025/26	2026/27	2027/28	2028/29	2029/30
Ofwat DD	19.5	18.3	17.2	16.0	14.8	13.7
TPI1	34.8	35.1	35.4	35.6	35.9	34.8
TPI2	28.0	28.3	28.1	28.4	28.7	28.0
Triangulated	31.4	31.7	31.7	32.0	32.3	31.4

Triangulated reflects the average of the values from TPI1 and TPI2.
Source: Oxera analysis.

Contact

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

oxera.com