Economic Insight

August 2023

Economic Insight Ltd

CRITIQUE OF OFWAT'S APPROACH TO INDICATIVE ODI RATES

Report for Welsh Water

CONTENTS

This report provides a critique of the indicative outcome delivery incentive (ODI) rates that Ofwat has calculated for companies to use in their PR24 business plans. We find that Ofwat's 'top-down' approach has material shortcomings, both in principle and in practice. We subsequently assess different options for how companies could calculate alternative ODI rates.

1 Introduction and executive summary

1A. Introduction

Ofwat's PR24 method directs and incentivises companies to adopt its own view of incentive rates for common PCs

As part of developing their business plans for PR24, companies must determine what ODI rates to apply to their performance commitments (PCs). Relatedly, Ofwat's final methodology for PR24 states that, to meet its 'minimum expectations' under its quality and ambition assessment (QAA), companies should adopt its proposed (standard) incentive rates for common PCs.[1] However, the regulator's method also states that companies can deviate from its proposed rates and still meet its minimum expectations, if they provide 'compelling evidence' as to why this is appropriate.[2]

Companies are also financially incentivised to adopt Ofwat's proposals under the QAA framework. Specifically, if by not adopting Ofwat's indicative incentive rates, a company's plan was graded as 'inadequate' by Ofwat, they may face a -30 basis points reduction in their allowed equity return, and less favourable cost sharing rates.[3] As such, Ofwat's method requires companies to carefully consider whether, and under what circumstances, it may be appropriate and necessary to deviate from the regulator's own proposed incentive rates. Broadly, companies must balance:

- their own objective view as to what incentive rates are in the best interests of customers, the environment, and other stakeholders; against
- the risk of financial penalty under any deviation from Ofwat's proposals (where those diverge from the former).

---

1 *'Creating tomorrow, together: Our final methodology for PR24 Appendix 12 – Quality and ambition assessment.' Ofwat (December 2022); page 7.*

2 *Furthermore, under Ofwat's method, it is not clear that a failure to comply with Ofwat's minimum expectations in relation to any one item in isolation (such as incentive rates for common PCs) necessarily results in a failure to meet Ofwat's quality assessment. This is because Ofwat describes its quality assessment as being 'in the round'.*

3 *'Creating tomorrow, together: Our final methodology for PR24.' Ofwat (December 2022); page 151.*

Subsequent to the publication of Ofwat's methodology, the above considerations have been made more challenging, due to relatively substantial changes in Ofwat's method for deriving its indicative standard incentive rates, as we outline below.

Ofwat determined that its original bottom-up method for standard incentive rates was not robust

In its final methodology, Ofwat originally stated that it would adopt a 'bottom-up' approach to setting standard ODI rates. The regulator explained that the aim of a bottom-up approach is to align the incentive rates to customer benefits (and costs) *relating to the specific PC in question.*[4] In practice, Ofwat proposed that standard incentive rates would be based only on estimated customer marginal benefits.[5] Ofwat explained that it considered a 'bottom-up' approach to be preferable to a 'top-down' approach to determining incentive rates: *"a bottom-up approach enables us to set incentive rates based on marginal benefits, which aligns the interests of companies with those of their customers and the environment, so that customers as a whole pay or receive ODI payments that broadly reflect the impact from a unit of outperformance or underperformance. This encourages companies to focus on what matters to customers."*[6]

Ofwat further sets out that, in relation to (the majority of) common PCs, its approach to measuring marginal benefits (and therefore, for setting the standard incentive rates) would be to base them on 'collaborative customer research' it was commissioning.[7] Ofwat explained that the benefit of this should be that it improved consistency across companies. However, once the results of the collaborative customer research were available, Ofwat considered that they were not appropriate to use for setting incentive rates alone. Ofwat's reservations related to challenges associated with interpreting the survey results as well as robustly mapping those results to PCs.

Ofwat therefore revised its approach to adopt a 'top-down' method

In light of the above concerns, on June 2nd Ofwat informed companies of its intention to amend its approach to setting standard incentive rates for common PCs.[8] Specifically, Ofwat's proposal is now to determine the rates using a top-down method. Unlike a bottom-up method (which, as above, sets incentive rates based on direct measures of customer benefits / costs for the PC in question), a top-down method sets individual incentive rates based on an amount of Return on Regulatory Equity (RoRE) at risk for each PC. More specifically, Ofwat's top-down method generally consists of:

---

4 *'Creating tomorrow, together: Our final methodology for PR24 Appendix 8 – Outcome delivery incentives.' Ofwat (December 2022); page 6.*

5 *This was due to concerns Ofwat had regarding the accuracy of marginal cost estimates, although Ofwat stated that it may take marginal cost evidence into account under its determinations, if it considered it sufficiently reliable (most likely as a cross check).*

6 *'Creating tomorrow, together: Our final methodology for PR24 Appendix 8 – Outcome delivery incentives.' Ofwat (December 2022); page 6.*

7 *'Creating tomorrow, together: Our final methodology for PR24 Appendix 8 – Outcome delivery incentives.' Ofwat (December 2022); page 13. Note, for biodiversity and operational GHG, Ofwat proposed to use external estimates.*

8 *Letter to ODI working group members from Ofwat, June 2nd 2023.*

- **Starting with an average RoRE allocation per PC of 0.5% RoRE**. Ofwat supports this choice with a historical analysis of company performance in the first two years of AMP7.

- **Selecting either 0.4%, 0.5% or 0.6% RoRE for each individual PC, based on customer research**. Ofwat allocates 'higher importance' PCs 0.6% RoRE, whilst 'lower importance' PCs are allocated 0.4% RoRE. Ofwat's choice is based on the outputs of three pieces of customer research.

- **Dividing the RoRE allocation across a performance range**. For each PC, for each company, for each available year, Ofwat calculates the historical deviation of actual performance from the PC level. For each PC, Ofwat calculates the P90 and P10 of historical deviations and selects the larger of the two measures. The PC's RoRE allocation is then effectively split across the RoRE range to derive a unit rate.

Ofwat's June 2nd letter to companies stated that its approach to the QAA framework was unchanged; and thus, to meet its minimum expectations, companies should adopt the indicative rates from its top-down method, or provide compelling evidence for not doing so. In addition, the letter asked for feedback from companies regarding Ofwat's top-down approach and the resultant incentive rates.

On June 15th, Welsh Water (Welsh) received its own indicative ODI rates from Ofwat.[9] A dataset with the indicative incentive rates for all companies was shared with companies on June 27th, and the calculations for all incentive rates was shared with companies on July 19th.[10]

We have developed a framework to assist Welsh in determining its approach to this issue

In the above context, Welsh commissioned Economic Insight to undertake a critique of Ofwat's updated approach to determining standard incentive rates for common PCs. The purpose of our work is to address three questions:

- In principle, when / under what circumstances would it be appropriate to deviate from Ofwat's proposed indicative rates, and why?

- In practice, where might this be the case at PR24?

- What alternatives should be considered, where deviation is appropriate?

This report is therefore structured around the above, whereby we firstly set out a framework for our critique (i.e. 'what' makes a good incentive rate?). We then apply our framework to Ofwat's current proposed method and rates. Finally, we outline proposed alternative approaches that Welsh could consider, were it to deviate.

[9] *On June 15th Ofwat shared Welsh Water's full set of rates, excluding river water quality. Rates for some PCs were shared in advance of this date.*

[10] *Ofwat shared its full suite of models on July 19th. A subset of models were shared on July 7th.*

1B. Executive summary

As summarised further below, we find that Ofwat's approach has material shortcomings, both in principle and in practice. We suggest that Welsh should not accept Ofwat's indicative ODI rates, and propose either that PR19 incentive rates are broadly retained or that Ofwat's top-down approach is adjusted.

Our framework for assessing Ofwat's approach

In order to assess Ofwat's approach to calculating indicative incentive rates (and subsequently evaluate alternative options), we have developed a set of criteria that a well-designed approach should meet. These criteria are as follows.

- **Criterion 1: The approach is, in principle, likely to give rise to appropriate incentives**. By this, we mean the approach gives incentive rates that result in companies being best-off when they provide the (economically) efficient service level for each outcome. This is important to ensure that companies focus their efforts and resources on the 'right' outcomes and make trade-offs between them that genuinely reflect costs and benefits to their customers (and wider society / the environment).

- **Criterion 2: The approach takes into account whether there are material company-specific differences in costs / benefits**. This is an extension to the first criterion. It recognises that the benefits and costs of improving performance levels across (efficient) companies will likely vary, e.g. depending on geography. Such differences should be reflected in incentive rates – such that each company is incentivised to focus on the right outcomes for their customers.

- **Criterion 3: The estimation method is robust and reliable**. This criterion relates to how robust and reliable the estimation method is. We would expect a well-designed method to be: based on reliable underlying data; with sufficient sample sizes; and to produce incentive rates that are not overly sensitive to changes in individual assumptions.

- **Criterion 4: The approach does not produce significantly different incentive rates from PR19**. This criterion reflects the importance of regulatory consistency over time, particularly in long-lived infrastructure industries. Significant changes in incentive rates over time are contrary to sound long-term planning and increase investor risk, which is detrimental to customers over time. Therefore, strong evidence should be required to materially change incentive rates between price controls.

- **Criterion 5: The approach is consistent with a balanced package of risk**. Finally, a well-designed approach should support a balanced package of risk for an efficient firm. This is important because, for a notional firm to be financeable (for equity) over PR24, it is necessary that investors' expected returns are equal to the allowed return on equity.

Summary of our assessment and findings

We find that Ofwat's approach fails against all of our criteria. Our principal concerns with Ofwat's approach are as follows.

- **Ofwat's approach is not designed to result in incentive rates that will encourage companies to focus their efforts in the right areas**. Ofwat's approach of selecting an amount of RoRE at risk for each PC, and then dividing that amount across a performance range will, conceptually, not result in incentive rates that encourage companies to improve performance up to the point at which the costs of improving further are more than the benefits that will be accrued. Ofwat's approach lacks a sound conceptual basis, is overly arbitrary, and contains internal inconsistencies. The approach therefore fails criterion 1. As a result, Ofwat's incentive rates can be expected to be either: (i) too small, meaning that a company does not have sufficient incentive to improve or maintain levels of service; or (ii) too high, meaning that a company may spend more on improving performance than would be in customers' interests.

- **Ofwat's application of its top-down approach lacks reliability and robustness**. Ofwat's approach therefore fails criterion 3. More specifically:

  - Ofwat's choice of the average amount of RoRE to allocate to each PC (0.5% RoRE) is arbitrary. Based on Ofwat's logic and the evidence it presented, it could have selected a materially different number.

  - The alignment between the customer research Ofwat has relied on and the definitions of PCs is weak. The amount of RoRE Ofwat has initially allocated to each PC is therefore highly subjective.

  - We have identified issues related to the performance range Ofwat has used to divide its selected RoRE at risk over, specifically in relation to: (i) unbalanced samples, whereby some companies contribute more observations and likely skew resulting performance ranges; (ii) small sample sizes, which means that performance ranges are unlikely to be reflective of the full range of likely performance; and (iii) Ofwat's distribution for demand PCs (leakage, PCC, and business demand) is calculated based on an aggregation, which is not based on sound logic and is inconsistent with Ofwat's broader approach.

- **Ofwat's approach does not provide consistency over time, and therefore damages long-term incentive properties**. Ofwat's indicative ODI rates for PR24 differ materially from those set at PR19 (once compared on a like-for-like basis[11]) – and there is no compelling evidence that Ofwat's indicative rates for PR24 are more appropriate than those at PR19. Such a change in incentive rates will both: (i) damage the credibility and incentive power of the PR24 rates; and (ii) more widely damage incentive power and increase regulatory risk, because investment decisions over PR19 will have been based on the PR19 rates and an expectation that they will not change significantly between price controls. Ofwat's approach therefore fails criterion 4. Ofwat's indicative rates would also likely increase the amount of RoRE that would be at risk from ODIs at PR24, and there is no evidence that this should be the case.

Further to the above, due to Ofwat's indicative rates being symmetrical, its proposed ODI package is likely to contribute to a skewed balance of risk from ODIs. It therefore also fails criterion 5. However, any skew is a function of both the ODI rates and where Ofwat sets the PC levels (the latter of which is out of scope of this work).

Given the above, we suggest that Welsh does not accept Ofwat's indicative incentive rates and instead proposes an alternative approach.

Alternative approaches

We have broadly considered four options:

- Ofwat's original bottom-up approach for PR24;
- Ofwat's top-down approach for PR24;
- An amended top-down approach; and
- Using PR19 incentive rates.

We have ruled out Ofwat's original bottom-up approach for PR24 for the same reasons as Ofwat has; and we rule out Ofwat's top-down approach for the reasons outlined above.

We suggest that Welsh adopts either of the other two approaches (or a combination of them).

[11] *Comparisons between PR19 and PR24 rates are made on the basis of a consistent price base and consistent normalisation parameters (e.g. number of connections).*

- The advantage of using the PR19 incentive rates is, being primarily derived bottom-up, they are (in principle) consistent with creating incentives for the company to prioritise its efforts in a way that will benefit its customers. They also avoid any material increase in investor risk that could further increase costs to customers over time. Welsh could adopt either its own PR19 incentive rates or the industry average – the latter having the potential for consistency across the industry.[12] Adjustments would need to be made to the PR19 incentive rates to account for inflation and growth in, for example, customer numbers / connections. An alternative approach would also be required for PCs that are new at PR24.

- The advantage of the amended top-down approach is that it achieves consistency across companies, whilst can still result in incentive rates for Welsh that are relatively consistent with those used at PR19. It can be used to target a total amount of RoRE at risk, and it can also provide for a greater degree of judgement in setting rates – although this should be exercised with caution. It can be used for both new and existing PCs.

- It could further be appropriate for Welsh to triangulate across the above two options, either 'across the board', or more selectively by PC (i.e. choosing whichever it felt was most suitable, in light of our criteria, PC-by-PC).

The rest of this report is set out as follows.

- Chapter 2 details our framework for critique.

- Chapter 3 details our assessment of Ofwat's approach against our framework.

- Chapter 4 presents our consideration of alternative approaches.

[12] *We note that at PR19, Welsh generally adopted a top-down approach in the incentive rates that it proposed. These were subsequently subjected to Ofwat's triangulation approach / application of a 'reasonable range'.*

2 Framework for critique

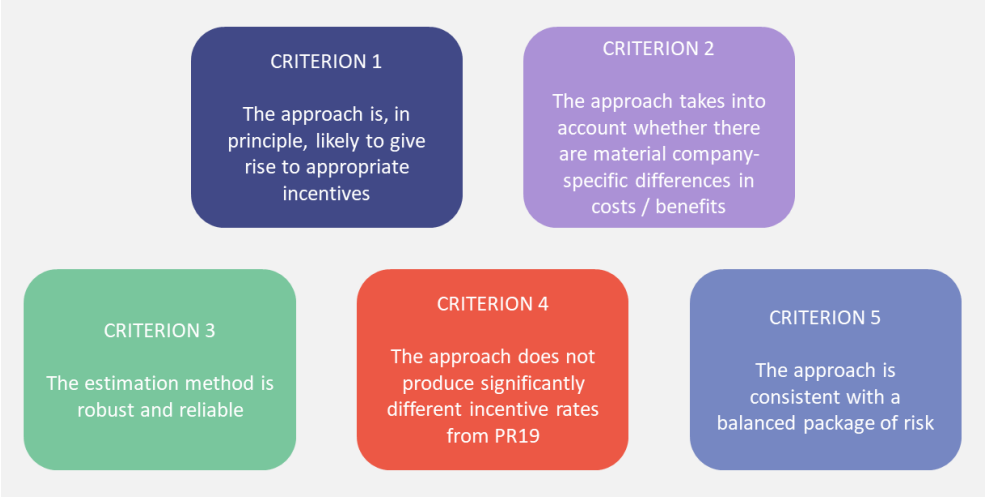
In order to assess Ofwat's top-down approach for deriving indicative incentive rates (and subsequently evaluate alternative options), we have developed a set of criteria that a well-designed approach should meet. In this chapter, we therefore detail our framework; first providing a brief overview of it, before then describing each criterion in turn.

2A. Overview of framework

Our evaluation framework is illustrated below in Figure 1. As can be seen, there are 5 criteria:

- the approach is, in principle, likely to give rise to appropriate incentives;
- the approach takes into accounts whether there are material differences in costs and benefits across companies;
- the estimation method is robust and reliable;
- the approach does not produce significantly different incentive rates from PR19; and
- the approach is consistent with a balanced package of risk (for an efficient firm).

Figure 1: Framework for assessing Ofwat's ODI rate approach



Source: Economic Insight

2B. Criterion 1: The approach is, in principle, likely to give rise to appropriate incentives

The first criterion is that the approach is likely, in principle, to give rise to appropriate incentives. By this, we mean the approach results in companies being best-off when they provide the (economically) efficient service level for each outcome. In other words, where the marginal cost (MC) of making further improvements is equal to the marginal benefit (MB) of achieving that improvement.

The importance of this criterion is that it ensures companies focus their efforts and resources on the 'right' outcomes and make trade-offs between them that genuinely reflect costs and benefits to their customers (and wider society / the environment). Indeed, *the entire purpose* of an outcomes framework is, fundamentally, to help achieve this end. Conversely, without any link to MB or MC, there is a real risk that an incentive regime would encourage companies to do things that are not in the best interests of customers, society, or the environment. For example, investing significant resource and effort (and incurring costs) to achieve a particular objective, where their customers would prefer them to do something else.

In practice, there are challenges with estimating MB and MC. Notwithstanding this, however, for any approach to produce ODIs that give rise to appropriate incentives 'in principle', it must factor in (*to some degree*) both: (i) the value that customers place on improving performance; alongside (ii) the costs of improving performance.

Ofwat has consistently recognised the above. For example, at PR19 the regulator set incentive rates using the following formulae:[13]

$$Underperformance\ rate = MB - MC * (1 - cost\ sharing\ rate)$$

$$Outperformance\ rate = MB * cost\ sharing\ rate$$

Similarly, in its final methodology for PR24, Ofwat's proposal was to set incentive rates based on its bottom-up approach that sought to directly measure MB because it: "*aligns the interests of companies with those of their customers and the environment*". Whilst Ofwat's method for PR24 did not include MC in the incentive rate formulae, as previously noted, the regulator's approach does nonetheless include a provision for it to take MC data and evidence into account in determining the rates (explaining that basing rates on marginal costs should also "*fund the costs of improvements and return that funding to customers for under-delivery*").[14]

[13] *'PR24 and beyond: a discussion paper on outcome delivery incentives', Ofwat (February 2022), page 7.*
[14] *'PR24 Final Methodology: Appendix 8 - Outcomes delivery incentives', Ofwat (December 2023), page 6.*

2C. Criterion 2: The approach takes into account whether there are material company-specific differences in costs / benefits

This criterion is an extension of our first. Put simply, it is logical to expect that (for some PCs) the MB and (efficient) MC of delivering specified levels of performance will differ across companies. Intuitively, there are reasons to suppose that differences in MC could be relatively material, due to differences in company supply areas (e.g. geography and topography) that affect the efficient costs of delivering certain infrastructure and the availability of raw water.[15] Differences in MB may be less pronounced (i.e. as a whole, although water customers are diverse, they may *broadly* care about the same things in a way that does not vary systematically by water company). Nonetheless, one would still not expect customer preferences to be exactly the same across all companies. Following from this, for an approach to create appropriate incentives (in principle), it must also allow for some degree of variation in incentive rates across firms.

In practice, the extent of genuine variation in MB or MC across companies is difficult to ascertain with precision. We note that one of the intended benefits of Ofwat's collaborative customer research at PR24 was that (whilst no method for estimating MB is perfect) a consistent approach would better allow the regulator and industry to distinguish between variation in MB that arose due to method variation, as opposed to it reflecting actual differences in customer preferences across companies. Whilst, in practice, the limitations with the collaborative research and results mean that Ofwat has deemed it not possible to use it in the way it originally intended, the underlying rationale for consistent research methods is sound. At this time, there is no one single source of robust, comparable, MB or MC data across companies.

Ofwat's final methodology for PR24 recognises the above, where it explained that it wanted to: "*set rates in a consistent way between companies, while allowing for material differences in customer preferences*".[16] Again, as previously noted, Ofwat's method also enables it to take MC evidence into account (where the regulator considers this suitably robust). Hence, the PR24 method allows for the possibility of varying incentive rates to reflect differences in the (efficient) MC between companies.

---

15 *In the same way that Ofwat's base cost models predict quite different 'efficient' costs across individual companies, given their differing characteristics.*

16 *'PR24 Final Methodology: Appendix 10 – Outcome delivery incentives, Ofwat (December 2023), page 7.*

2D. Criterion 3: The estimation method is robust and reliable

This criterion relates to how robust and reliable the estimation method is, as used under any approach to setting incentive rates. To be robust and reliable, we mean that:

- The underlying source / method for any data and evidence used to infer the incentive rates is reliable. For example, where customer research is the source for MB, the research method used is in line with best practice; is appropriately implemented; and gives plausible results.
- The data sample (evidence base) used in the calculation of incentive rates is sufficiently large and complete.
- The calculation method used to derive the incentive rates is robust and the rates are not highly sensitive to changes in individual assumptions.

Any approach that fails to sufficiently meet the above requirements would risk creating poor incentives, which may be to the detriment of customers, the environment and other stakeholders. Indeed, Ofwat itself has acknowledged the importance of this in its decision to move away from its proposed 'bottom-up' approach at PR24, in light of concerns regarding the collaborative research results.

In applying this criterion in practice, we have focused on the second two of the above bullets. This is because a full review of the respective merits of prior underlying customer research is outside the scope of our work for Welsh.

2E. Criterion 4: The approach does not produce significantly different incentive rates from PR19

This fourth criterion reflects the importance of regulatory consistency over time, particularly in long-lived infrastructure industries. Put simply, if incentives change 'drastically' from one price control to the next, the risk is that companies are repeatedly tasked with 'changing direction'; moving their focus from one goal to another, with little warning, every five years. Even if that were appropriate from a short-term efficiency point of view, it is contrary to sound long-term planning and increases investor risk, which is detrimental to customers over time.

This criterion must also be considered in the context of the previous one regarding the robustness and reliability of any estimated incentive rates. For example, suppose a revised method gave rise to a very different incentive rate than had applied at a previous price control. Given the inherent challenges in measuring MB and MC, one would want to take considerable care to ensure that said 'new method' was, in fact, an improvement that better reflected benefits and costs, as opposed to the difference simply reflecting a 'method change' that was (at best) neither objectively more, nor less, reliable than prior estimates. Furthermore, even if one was confident that a 'new method' gave rise to an incentive rate that better reflected MB and MC (compared to a previous price control), a significant change in said rate remains problematic. This is

because companies (and investors) will have made decisions that reflected the prior incentive rates. Radical changes in rates thus change the risk-reward balance for those decisions ex-post, which (all else equal) may increase the cost of capital.

Of course, the incentive rates used at PR19 are themselves subject to measurement error (indeed, Ofwat raised concerns regarding the extent of variation in rates across companies at that time). They do not (nor could any method ever) perfectly measure the 'true' MB and MC to customers of companies' service improvements. We do not, therefore, advocate anchoring the industry to those estimates for consistency reasons alone indefinitely. Rather, the appropriate solution (beyond PR24) is to derive more robust estimates of both MB and MC across the industry. However, at this time, there is neither:

- a revised method / evidence source for MB and MC across the industry that one could confidently say is objectively superior to those available at PR19; nor
- any a priori reason to believe that MB and MC will have drastically changed since PR19.[17]

We have also considered reasons why incentive rates might legitimately change over time. In particular:

- **Inflation** will mean that MB and MC are higher in nominal terms, and therefore one would expect nominal incentive rates to increase between successive price controls – but this would not cause real incentive rates to change.
- **Cost sharing rates** will affect the marginal cost that both companies and customers incur, and (as can be seen by the incentive rate formulas at PR19)[18] will affect the optimal incentive rates set for the companies. Although Ofwat has changed some of the aspects of cost sharing rates at PR24 compared to PR19, we do not consider that this should result in different incentive rates at PR24. Notably: at PR19, the three fast track companies received cost sharing rates of 50%, and other companies generally received a rate based on how their submitted costs compared to Ofwat's view of costs;[19] and at PR24, companies ranked 'outstanding' and 'standard' will receive a 50% cost sharing rate, with others getting a higher rate on overspend and a lower rate on underspend.[20]
- We would not expect **increasing RCVs / Regulatory Equity** to necessarily lead to an increase in appropriate incentive rates for individual PCs. More specifically:

[17] *MC could be expected to increase because, when higher performance levels are reached, it gets incrementally harder (more costly) to increase performance further.*
[18] *See section 2B.*
[19] *PR19 cost sharing rates are set out in 'PR19 Final Determinations: Securing cost efficiency technical appendix', Ofwat (December 2019), table 24.*
[20] *'Creating tomorrow, together: Our final methodology for PR24: Appendix 9 Setting expenditure allowances', Ofwat (December 2022), Table 2.3.*

- One reason for RCV growth over the PR24 period is the investment required to meet new environmental standards. The related assets have no bearing on, for example, the benefits that customers derive from reducing supply interruptions. It is not clear that they would have a bearing on the costs of supply interruptions either. As such, RCV growth driven by environmental standards should have no bearing on appropriate incentive rates for supply interruptions.

- Another reason why there might be increasing Regulatory Equity is because Ofwat has reduced its notional gearing assumption from 60% at PR19 to 55% at PR24. Ofwat's notional gearing assumption does not directly affect the MB or MC of improvements in performance levels, and therefore should not lead to higher incentive rates.

- Set against the above, RCV growth may in part be driven by an increasing number of connections and other 'scale factors'. For example, given the units of measurement for some incentive rates (e.g. £ per incidents per 10,000 connections), a higher number of connections would lead to a higher incentive rate, all else equal. This is because, for example, a company with twice as many connections would require an incentive rate that is twice as high, to reflect the same marginal costs and marginal benefits on a per incident basis. We understand that the growth in connections and other scale factors is relatively low and therefore should not result in significant increases in incentive rates.

Given the above, we consider that an appropriate approach to PR24 should be one that results in a reasonable degree of consistency with the rates used at PR19.

Going forward, if a method for PR29 is developed that addresses some of the substantive issues identified at PR24 (particularly with regard to the collaborative customer research), one could then revisit the issue of consistency. At that time, if the new (objectively better) method gave quite different estimates than the past, a glide path or similar could be considered, so as not to create undue uncertainty and customer harm. Furthermore, any robust method that was used at PR29 should logically be repeatable in future controls in a way that gives rise to broadly similar estimates (i.e. because, intuitively, MB and MC should not be radically changing over relatively short periods of time).

2F. Criterion 5: The approach is consistent with a balanced package of risk

The final criterion is that the approach to incentive rates should be consistent with a symmetrical balance of risk (for an efficient firm). This is important because, for a notional firm to be financeable (for equity) over PR24, it is necessary that investors' expected returns are equal to the allowed return on equity (i.e. all else equal, investors expect neither to out-, nor under-, perform against regulatory incentives).

First, focusing specifically on incentive rates (as opposed to other aspects of the ODI framework), an approach can be consistent with a balanced package of risk if: (i) there are both rewards and penalties for each PC; and (ii) the incentive rates for rewards and penalties are the same (i.e. are symmetrical). In addition to the above, the amount of RoRE risk associated with each PC (i.e. therefore, the absolute size of the unit incentive rates) can also affect the risk balance for companies at PR24. For example, because efficient firms face trade-offs, a notionally efficient firm may outperform on some PCs; underperform on others; but *overall* will expect to neither out-, nor under-, perform. So, suppose that at PR24, the incentive rate method allocated 'more RoRE' to PCs where underperformance was more likely (or less RoRE where outperformance was more likely) as compared to PR19. In that case, even with symmetrical incentive *rates*, the incentive rate method itself may not result in a symmetrical risk balance.

Further to the setting of incentive rates, it is also important to recognise that, in practice, other elements of the ODI framework may not be consistent with a balanced risk package (for an efficient firm). For example, it currently remains the case that performance commitment levels (PCLs) are set based on extrapolations, rather than being based on the efficient level. In addition, there are elements of the regulatory framework that are penalty only (such as PCs related to statutory compliance).[21] In that context, to achieve overall symmetry, one *could* consider offsetting any skew arising elsewhere by having asymmetric incentive rates. We would, however, advocate addressing any asymmetry 'at source' (i.e. appropriately calibrating PCLs) rather than using incentive rates as a counterbalance.

The final consideration relating to the risk balance is the 'overall' amount of RoRE risk allocated to ODIs, as faced by companies. Large changes in this (relative to the risk allocated to other elements of the framework) may affect both the overall risk-reward balance and the relative incentives for companies to focus on service delivery, as opposed to cost efficiency. This would be contrary to regulatory consistency, and thus may increase investor risk (and thus, costs to customers) in the long-term. Moreover, 'excessive' or 'inappropriate' risk exposure under the regulatory framework (e.g. allocation of risk to companies that is outside of management control) could undermine the ability of efficient firms to be financially resilient or financeable.

For PR24, Ofwat's view of what 'extent' of risk an efficient company should face from ODIs is around ±1% to ±3% RoRE each year.[22, 23] Ofwat had a similar view at PR19.[24] We would therefore not expect to see material changes in the extent of risk that companies are subject to between price controls, under an appropriate method for setting incentive rates.

---

[21] *'Creating tomorrow, together: Our final methodology for PR24 Appendix 8 – Outcome delivery incentives.' Ofwat (December 2022); page 12.*

[22] *'PR24 Final Methodology: Appendix 10 – Aligning risk and return', Ofwat (December 2023), page 15.*

[23] *We note that Ofwat has not put forward any evidential basis for this risk range.*

[24] *'PR19 Final Determinations – Aligning risk and return technical appendix', Ofwat (December 2019), page 27.*

3 Evaluation of Ofwat's approach

This chapter details our assessment of Ofwat's approach, based on the framework set out in the previous chapter. For each criterion, we first set out the relevant parts of Ofwat's approach, and then provide our assessment of the extent to which the criterion is met. We summarise our assessment of each criterion with a red-amber-green (RAG) rating.

3A. Criterion 1: The approach is, in principle, likely to give rise to appropriate incentives

Ofwat's approach

Ofwat's overall approach involves the following two key steps:

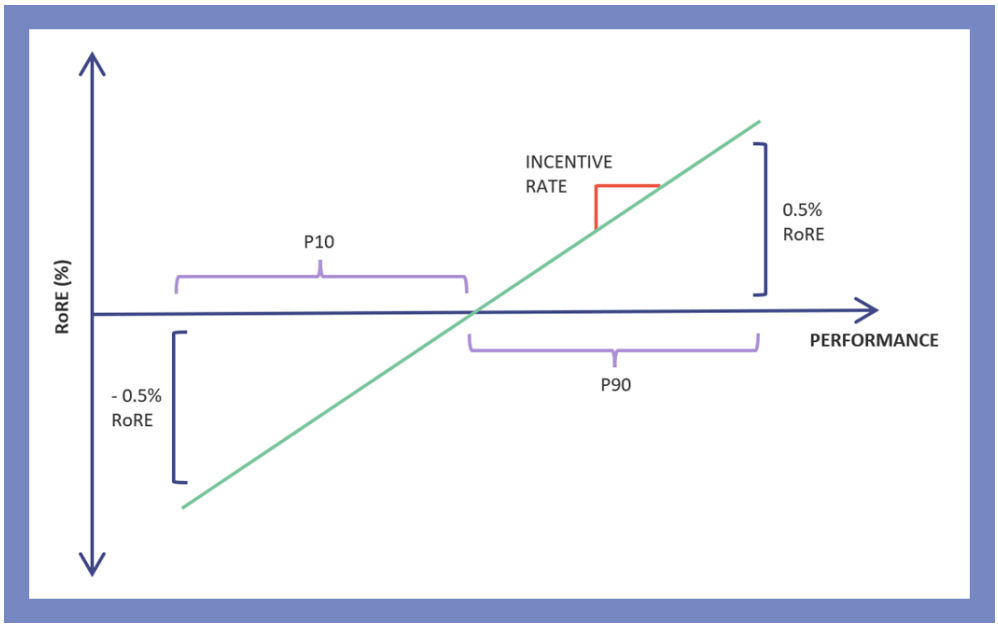
(i) Each PC is first allocated a "*theoretical maximum*" percentage of RoRE at risk.[25]

(ii) The RoRE at risk per PC is then divided over a "*stretching but achievable*" performance range, to arrive at a unit incentive rate.[26]

This is illustrated in the following stylised figure.[27]

[25] *'Top-down approach: Using customer preferences to inform ODI rates', Ofwat (June 2023), page 3.*
[26] *'Top-down approach: Overview', Ofwat (June 2023), page 1.*
[27] *This figure is a stylised representation of Ofwat's approach. We acknowledge that there are a number of steps to Ofwat's calculations, to arrive at a standardised unit incentive rates expressed in £s, that are not captured in this figure.*

Figure 2: Stylised illustration of the key elements of Ofwat's approach



Source: Economic Insight

With regard to (i), Ofwat selects a RoRE allocation of either 0.4%, 0.5% or 0.6% for each PC. There are two key elements to this step.

- First, Ofwat seeks to set the average % RoRE at risk per PC as the theoretical maximum amount of risk a company is exposed to. Ofwat selected 0.5% RoRE, and supported this with an analysis of hypothetical historical ODI payments over the first two years of PR19 (2020-21 and 2021-22). More specifically:
  - For each PC, Ofwat calculated hypothetical ODI payments by multiplying deviations from PCLs by the underperformance rate (excluding the effect of any caps and collars). Ofwat then took the UQ and P90 of these payments in RoRE terms (treating all payments in absolute terms i.e. treating outperformance and underperformance equally).
  - The UQs and P90s are then averaged across all PCs. The average P90 hypothetical payment was shown to be 0.70% and 0.74% for water and wastewater PCs respectively; and the UQ payment was shown to be 0.35% and 0.25% for water and wastewater PCs respectively.[28] Ofwat selected 0.5%, which falls between these ranges, as the average allocation per PC.

[28] *'Top-down models – RoRE allocation.xls', Ofwat (31 July 2023). We note that the analysis presented in Ofwat's June note ('Top-down approach: Using customer preferences to inform ODI rates') contained a number of errors and was therefore subsequently corrected.*

- Second, the 0.5% allocation is then varied across the PCs, to account for relative differences in the importance customers place on the PCs. Specifically, 'lower' importance PCs are allocated 0.4% RoRE, 'medium' importance PCs are allocated 0.5% RoRE, and 'higher' importance PCs are allocated 0.6% RoRE. The variance in importance for customers has been established through triangulating 3 pieces of customer research. The % RoRE is finally calculated as a £s value for each company, using their regulated equity.[29]

With regard to (ii), Ofwat uses historical analysis to select the performance range per PC over which to divide the % RoRE. Key elements of this analysis are as follows:

- In general, Ofwat calculates the historical difference between actual performance and the PCL (expressed as a percentage), for each company and each year of available data. This data is then pooled across companies and years, providing a distribution of under- and over-performance for each PC.[30, 31]

- The P10 and P90 of this distribution are then taken. Whichever is the largest in absolute terms is then used as the performance range to calculate both the out- and under-performance rate, thereby giving symmetrical rates. Specifically, the performance range calculated is split across each company-specific £ RoRE value, to obtain a £s unit rate for each company. The median unit rate is then taken as the ODI incentive rate, thereby giving a consistent rate across all firms.

Our assessment

We have found that Ofwat's approach cannot be expected to give rise to appropriate incentives. This is for a number of reasons, as set out below. We also recognise a positive feature of Ofwat's approach at the end of this section.

Ofwat's approach is not designed to result in incentive rates that will encourage companies to focus their efforts in the right areas

Ofwat's approach is not based on either the marginal benefit that customers receive from improved performance, nor the marginal cost that they would face for such improved performance. Therefore, as a starting point, one cannot expect that it would provide appropriate incentives.

In its discussion paper on outcome delivery incentives published in February 2022, Ofwat itself acknowledges the key downfalls of using a top-down approach; namely that, due to the disconnect from the calculation of marginal benefits, "*it is possible that the resulting ODI rates do not reflect impacts on customers*".[32]

A top-down approach could produce incentive rates broadly in line with PR19, but Ofwat's methodological choices mean that, in practice, its rates are inflated

There are certain features of Ofwat's approach that suggest that it could give results that are broadly in line with rates from PR19. That is, Ofwat takes the hypothetical

[29] *Ofwat used 2022-23 RCV to calculate £ values from the % RoRE figures.*
[30] *'Top-down approach: Overview', Ofwat (June 2023), pages 1-2.*
[31] *Ofwat's distributions are illustrated in the annex to this report.*
[32] *'PR24 and beyond: a discussion paper on outcome delivery incentives', Ofwat (February 2022), page 17.*

UQ/P90 RoRE at risk over the first two years of PR19, and spreads it across a historical P90 performance range (generally from AMP7 and AMP6). Given the PR19 rates were generally based on bottom-up calculations that take both the marginal cost and marginal benefit into account,[33] one could expect Ofwat's top-down approach to give rates for PR24 that are broadly reflective of marginal costs and benefits.

However, as shown subsequently in Section 3D, Ofwat's indicative rates are often multiples of the PR19 rates. Reasons for this relate to Ofwat's methodological choices, such as:

- Ofwat takes the % deviation from the PCL observed historically (from AMP6 and AMP7), and applies this to the 2024-25 PCL to obtain the performance range over which the RoRE is split, expressed in levels. The PCL at 2024-25 will be more stretching (i.e. generally a smaller number) than the historical PCL, which means that the RoRE amount is spread over a smaller performance range than was in fact observed historically. Spreading a historical amount of RoRE over a smaller performance range will result in higher incentive rates.

- By selecting the % RoRE to be put at risk per PC based on an analysis of the UQ and P90 payments over 2020-22 (or in other words, the upper end of the distribution), Ofwat may be using companies that had higher incentive rates relative to others at PR19 to set the incentive rates for all companies at PR24. This will calibrate incentives with companies that had higher than average rates at PR19.

Ofwat's approach uses customer research to inform incentive rates for non-customer facing PCs

As Ofwat itself acknowledged, "*a top-down approach, even aligned to customers' relative priorities, may not adequately reflect the impact on customers from a service failure if the performance range is too narrow or too wide. Customers may struggle to effectively allocate their total willingness to pay for service incentives to individual service areas without a clear understanding of the impact on service performance*."[34]

This means that, as commonly understood, customer research will likely be less informative regarding the relative importance of measures that customers have less direct experience of, such as the asset health measures. Despite this, Ofwat uses customer research to allocate an amount of RoRE for leakage, PCC and business demand – measures of performance that customers do not have an intrinsic value for.

As a result of this, Ofwat's approach cannot be expected to provide appropriate incentives for non-customer-facing PCs. We discuss Ofwat's approach to sewer collapses, mains repairs and unplanned outages in section 3C.

[33] *Please refer to Section 2B.*
[34] *'PR24 and beyond: a discussion paper on outcome delivery incentives', Ofwat (February 2022), page 17.*

There are a number of internal inconsistencies in Ofwat's approach

Whilst Ofwat has not detailed the rationale behind its approach, these apparent inconsistencies further call into question the in-principle merits of Ofwat's approach. These inconsistencies include:

- In deciding the amount of RoRE that it allocates per PC, Ofwat states that it is targeting a "*theoretical maximum amount of risk a company is exposed to through ODI payments*".[35] Not only does this lack any logical basis, it is also inconsistent with Ofwat's choice of: (i) the starting 0.5% RoRE, which is part way between Ofwat's calculated UQ and P90 of hypothetical historical RoRE payments per PC; and (ii) a P90 level of performance deviation from PCLs. It is inconsistent because one would expect the maximum of distributions to be included in calculations pertaining to a 'theoretical maximum' risk. Ofwat's approach is therefore arbitrary.

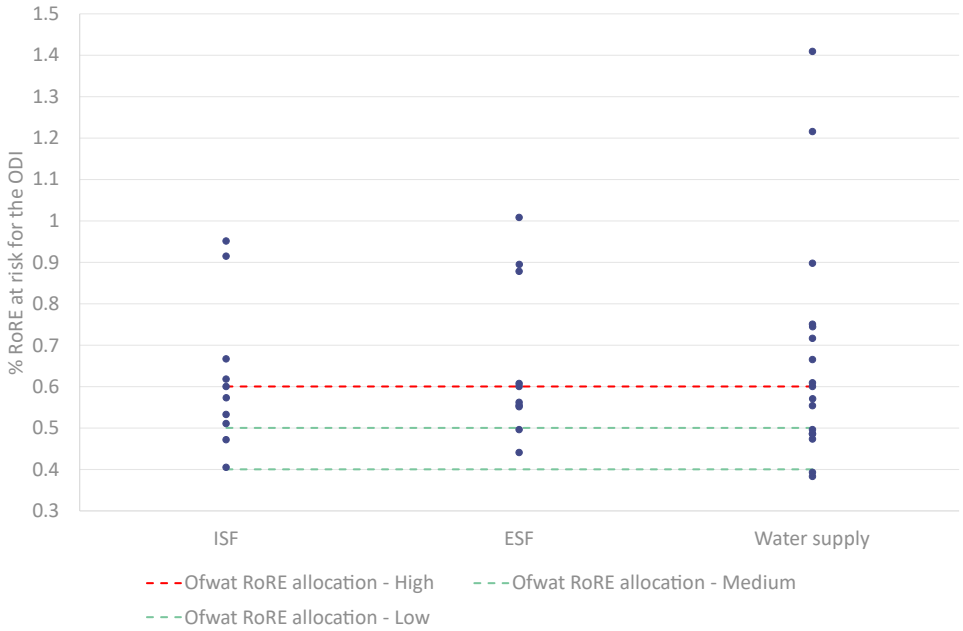
- Whereas Ofwat appears to consider it appropriate that each PC is allocated an amount of RoRE between 0.4% and 0.6%, that is not the amount of RoRE that each company actually has at risk (based on Ofwat's risk ranges). This is because, while Ofwat starts with 0.4% to 0.6%, it ultimately takes the unit rate for the **median firm**, to ensure a consistent rate across the industry. Ensuring a consistent % of RoRE is at risk across the industry and selecting a consistent unit rate across the industry are mutually exclusive. By ultimately prioritising the latter, the % RoRE at risk for each company will differ.

The final point above is illustrated in Figure 3 below. The figure shows how the RoRE at risk for three 'high importance' PCs varies significantly for each company compared to Ofwat's 0.6% target (assuming Ofwat's P90 that it uses to set the rates). Whilst the median firm will have 0.6% RoRE at risk, all other firms will be exposed to a different level of risk. The amount of RoRE that other companies will be exposed to depends on whether they have a higher or lower ratio of connected properties to regulatory equity, compared to the median firm.

Consequently, we find that whilst Ofwat's apparent target RoRE at risk is 0.6%, the risk companies are exposed to varies from under 0.4% to over 3.0%. This means that companies are exposed to significantly different risk to the level that Ofwat's approach initially intends, and for a number that is far higher.

[35] *'Top-down approach: Using customer preferences to inform ODI rates', Ofwat (June 2023), page 3.*

Figure 3: The 'actual' % RoRE Ofwat uses in its calculations (per PC for each company)

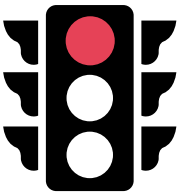


Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'. Note HDD (over 3% RoRE at risk) is excluded.

Positive in-principle features of Ofwat's approach

Despite the above flaws in Ofwat's approach, we nevertheless recognise that it does have a positive in-principle feature in that, all else equal, PCs that customers rank as more important will have higher incentive rates. This is consistent with PCs that have higher marginal benefits having higher incentives rates – as one would expect from a theoretical basis.

Conclusion



Overall, Ofwat's approach does not meet this criteria. Its approach is not designed to calculate the incentive rates that will encourage companies to focus their efforts in the right areas. Instead, Ofwat has developed an arbitrary approach that can be expected to result in incentive rates that are either:

- (i) too small, meaning that a company does not have sufficient incentive to improve or maintain levels of service; or
- (ii) too high, meaning that a company may spend more on improving performance than would be in customers' interests.

We therefore rate Ofwat's approach red in relation to this criterion.

Nevertheless, we also recognise the practical challenges in adopting a bottom-up approach that calculates incentive rates based directly on estimates of marginal benefits and marginal costs (or an approach that accurately infers appropriate incentive rates). We discuss this further in chapter 4 in relation to alternative approaches companies could adopt.

3B. Criterion 2: The approach takes into account whether there are material company-specific differences in costs / benefits

Ofwat's approach

For each PC, Ofwat's approach results in the same standardised incentive rate for all companies. For example, Ofwat's incentive rate for internal sewer flooding is identical across companies on a per incident basis (it differs on a per unit basis of the PC, which is incidents per 10,000 connections). Nevertheless, Ofwat has said that it will consider company-specific proposals if compelling evidence is provided.

Our assessment

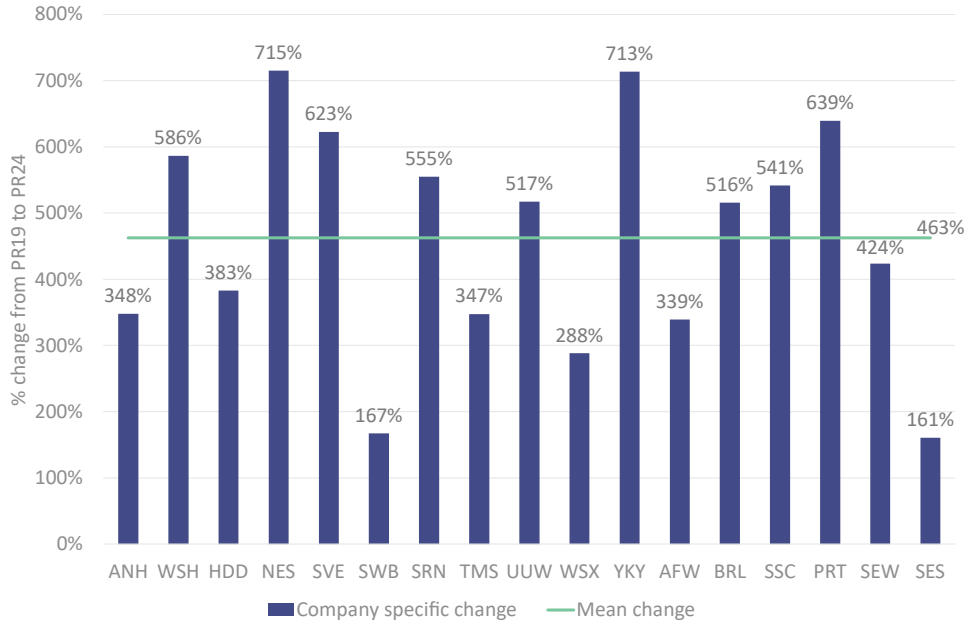
Ofwat's approach results in the same standardised incentive rate for each company. It has not presented any empirical analysis to show that this should be the case. On the basis that one would expect some variation between companies in terms of marginal benefits and marginal costs, we consider that Ofwat's approach therefore fails this criterion. We nevertheless recognise that: (i) Ofwat's approach allows companies to put forward evidence in relation to company-specific incentive rates; and (ii) in general, one would not expect large differences across companies.

We discuss the issues with Ofwat's approach in the remainder of this section.

Ofwat does not acknowledge that its PR24 rates vary significantly from its PR19 rates on a company-specific basis

Analysis of Ofwat's PR24 rates compared to PR19 rates shows that they vary by a greater degree for some companies than others. For example, the figure below shows that the industry average increase in rates for PCC is 463%; whereas the highest increase is 715% and the lowest increase is 161%. Although Ofwat expressed concern about variability between companies at PR19, it has not presented any evidence why such deviations were appropriate at PR19 but are now not at PR24.

Figure 4: Change in the PCC ODI rate for each company between PR19 and PR24



Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models' and 'Ofwat Final Determinations'

Ofwat does not recognise that customer preferences could vary between regions

By standardising the unit rate between companies, Ofwat assumes that customers in all regions place the same value on performance for each PC. In effect, this is assuming that all customers have the same marginal benefit for each PC.

However, this may not be the case. Customers might value performance on certain PCs differently between regions. Indeed, whilst Ofwat's own research data states that "*Welsh Water consumers consider service aspects that relate to water supply and environment in a similar fashion to English water consumers*", it found some attitudes that were different.[36]

- Welsh customer had a "*strong desire"* for reinvestment.
- Hose pipe bans "*lacked any relevance"* because customers felt that their regions received sufficient rainfall. This meant that they considered themselves to be at lower risk of drought.
- Bathing water quality was of "*some importance"* because the press had covered a recent drop in quality.

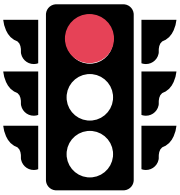
By standardising the unit rate in this scenario, Ofwat's incentive rates may be 'too high' in some regions and 'too low' in others.

As we have explained in Section 2B, setting incentives at a level that is not reflective of marginal benefit leads to suboptimal outcomes for customers. Setting marginal benefit

[36] *Yonder research, page 23.*

too high would incentivise companies to produce a higher quality product than what customers want.

Conclusion



Overall, Ofwat's approach does not meet this criterion. Its approach does not account for any company-specific differences. Failing to make company-specific adjustments means that incentives are unlikely to be a true reflection of the marginal benefit each region's customers receive, and the marginal costs to deliver performance. This will result in companies being incentivised to deliver suboptimal outcomes. We therefore rate Ofwat's approach red in relation to this criterion.

3C. Criterion 3: The estimation method is robust and reliable

Ofwat's approach

Below, we set out relevant parts of Ofwat's approach in relation to this criterion for: its choice of starting RoRE; its choice of allocation of PCs to a starting RoRE; and its choice of performance range for each PC.

Ofwat's choice of starting RoRE

As outlined in section 3A, Ofwat selects 0.5% RoRE as its average starting RoRE allocation per PC. This is based on an analysis of historical hypothetical ODI payments over the first two years of AMP7. Ofwat's choice of 0.5% is within the range of the historical UQ and P90 (as summarised in the table below).

Table 1: Ofwat's calculated hypothetical ODI payments, 2020-21 and 2021-22, % RoRE

	Historical UQ hypothetical ODI payment per common PC	Historical P90 hypothetical ODI payment per common PC
Common water PCs	0.35%	0.70%
Common wastewater PCs	0.25%	0.74%

Source: 'Top-down models – RoRE allocation.xls' (Ofwat)

Ofwat's allocation of PCs to a starting RoRE

Ofwat then varies the starting allocation to each PC between 0.4%, 0.5% and 0.6% RoRE for PCs it considers of higher, medium, and lower importance respectively. Ofwat's choice as to the importance of each PC is based on three pieces of customer research, as summarised in the figure below.

Figure 5: Ofwat's choice of PC importance

PC	Customer research 1 (ranking of valuations)	Customer research 2	Customer research 3 (% of customers ranking as high importance)	Overall proposed ranking
Internal sewer flooding	1 (H)	H	84% (H)	**H**
External sewer flooding	2 (H)	H	84% (H)	**H**
Water supply interruptions	3 (H)	H	83% (H)	**H**
Compliance risk index (CRI)	4 (H)	H	87% (H)	**H**
Customer contacts	5 (M)	H	87% (H)	**H**
Discharge permit compliance	N/A	M	82% (M)	**M**
Serious pollution incidents	6 (M)	M	82% (M)	**M**
Storm overflows	8 (M)	L	82% (M)	**M**
Total pollution incidents	7 (M)	M	82% (M)	**M**
River water quality	9 (L)	M	82% (M)	**M**
Leakage	N/A	M	81% (L)	**M**
Per capita consumption	N/A	L	79% (L)	**L**
Business demand	N/A	L	79% (L)	**L**
Bathing water quality	10 (L)	L	82% (M)	**L**

Source: 'Top-down approach: Using customer preferences to inform ODI rates' (Ofwat)

Ofwat's choice of performance range for each PC

As also outlined in section 3A, Ofwat's general approach to choosing the performance range for each PC is to:

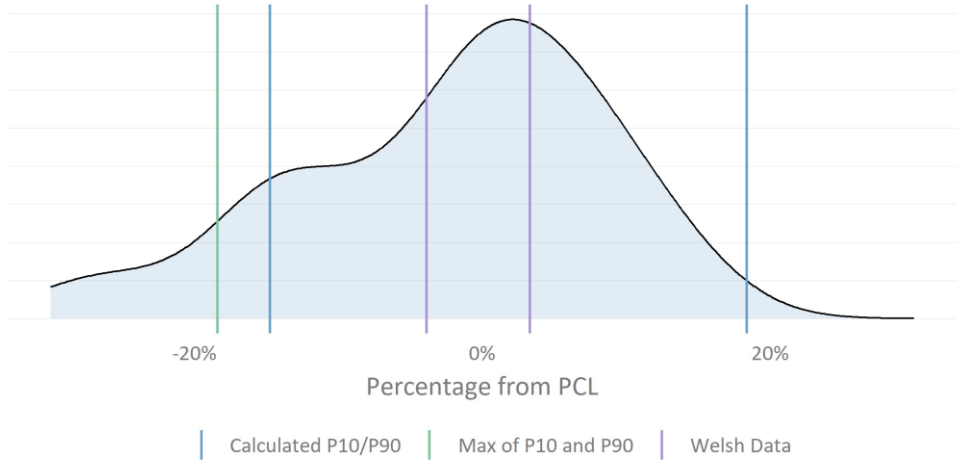
- calculate the historical difference between actual performance and the PCL in percentage terms, for each company, for each year of available data;
- pool these observations across companies and years, providing a distribution of under- and over-performance; and
- then select the larger of the P90 and P10 of the distribution.

The figure below illustrates Ofwat's approach for external sewer flooding.[37] In particular, it shows a smoothed distribution based on the 18 observations (9 companies, 2 years). Observations for Welsh are shown by purple lines; the actual P10 and P90 levels on the distribution are shown by blue lines; and the P10 (the maximum of the P10 and P90) applied to the opposite side of the distribution is depicted by the

[37] *The distributions for all PCs are shown in the annex to this report.*

green line. Ofwat chooses the larger of the P10 and P90, which in this case is an 18% deviation from the PCL.

Figure 6: Illustration of Ofwat's choice of performance range (external sewer flooding)



Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'.

There are a number of exceptions to Ofwat's general approach outlined above. In relation to **demand PCs** (PCC, business demand, and leakage), Ofwat calculates its performance range using aggregate demand data. It estimates this by summing performance data for PCC, business demand and leakage. The performance range for each of these three PCs is then based upon this aggregate demand distribution, rather than the performance data for each individual PC.

Ofwat has released a note explaining its decision.[38] Its states that Covid shifted PCC and business demand performance levels away from their historical levels. As Ofwat believes that PCC and business demand move in opposite directions away from their historical levels, aggregating the performance levels 'offsets' the impact.

In relation to **asset health** PCs (mains repairs, sewer collapses, unplanned outages), Ofwat calculates three ranges:

- (i) the **P90 of historical absolute percentage deviations from PCLs**, as per Ofwat's general approach;
- (ii) the **natural limit**, which is the difference between the PCL and the natural limit of performance, which is 0 (i.e. a performance range of 100%); and
- (iii) a **'reasonable performance'** based on Ofwat's judgement. Ofwat determines whether, in its view, the PR19 rates were effective. It then selects a 'reasonable performance' range that will adjust PR24 rates up or down based upon PR19 rates to make them more effective.

To derive its 'reasonable performance' range, Ofwat first calculates the median incentive rates using performance ranges (i) and (ii). It then compares the rates that these performance ranges produce to the equivalent median rates at PR19. Ofwat

[38] *'Top-down approach: Detailed note on demand indicative ODI rates calculation', Ofwat (2023), page 1.*

determines whether the PR19 incentives were effective at producing the right level of performance. Based upon this judgement, Ofwat decides whether the PR24 rates produced with performance ranges (i) and (ii) are likely to produce effective rates – or if they need to be adjusted up or down. Where an adjustment in the rates is required, Ofwat selects a 'reasonable performance' range that will generate the PR24 rate it is targeting.

In relation to **serious pollution** incidents, Ofwat takes the following steps to calculate an ODI.

(i) Performance for the PC is measured as the number of serious pollution incidents that occur annually. This is so that the PC is measured against the same units as the raw data is recorded in.

(ii) To set a rate that is comparable across companies, Ofwat transforms this performance data to the number of pollution incidents per 10,000km sewer network. By normalising by 10,000km sewer network, companies can be compared – it is reasonable to expect companies to have the same number of incidents over a standardised network length.

(iii) Ofwat sets a PCL and estimates a P10 / P90 range using this normalised data on the number of pollution incidents per 10,000km sewer network. Ofwat estimates a normalised performance range between the P10 / P90 and PCL.

(iv) For each company, Ofwat converts this performance range to the number of serious pollution incidents that a company is allowed, based upon its sewer network length.

(v) For each company, Ofwat then divides its RoRE at risk by the number of serious pollution incidents it is allowed to estimate a unit rate.

In effect, steps (iii) and (v) follow the same methodology that Ofwat has taken for all other PCs, but steps (ii) and (iv) differ. Whilst Ofwat has not given a reason for taking a different approach to other PCs, we consider that these steps are likely taken so a performance distribution can be compared across companies.

In relation to **CRI** and **discharge permit compliance**, these PCs are measured as a numerical score out of 100. Therefore, rather than estimating their performance range as a percentage (as it has done for all other PCs), Ofwat measures the performance range in terms of the numerical score. Similar to pollution incidents, this does not represent a change in methodology – as the P90 range is still being calculated the same way.

Our assessment

We set out our assessment for each of the three aspects of Ofwat's approach outlined above.

Ofwat's choice of starting RoRE

As identified in section 3A, Ofwat's choice of the starting RoRE lacks a conceptual basis and is therefore arbitrary. It has presented analysis that shows RoRE figures of a similar order of magnitude, but there is no conceptual reason for Ofwat to have picked a number between the P90 and UQ of hypothetical historical payments. Therefore, the question of whether Ofwat's choice of 0.5% RoRE is robust and reliable is somewhat limited.

The arbitrary nature of Ofwat's choice is demonstrated by the fact that: (i) Ofwat initially supported its choice of 0.5% with analysis presented in its June note;[39] and (ii) when it revised its analysis due to errors, which materially reduced its calculated ranges, its choice of 0.5% did not change. The table below shows Ofwat's original and revised ranges.

Table 2: Ofwat's original and corrected analysis to support its choice of average starting RoRE

		Average UQ payment per PC	Average P90 payment per PC	Mid-point between UQ and P90
Ofwat's original analysis	**Water PCs**	0.36%	0.77%	0.57%
	Wastewater PCs	0.47%	1.05%	0.76%
Ofwat's corrected analysis	**Water PCs**	0.35%	0.70%	0.53%
	Wastewater PCs	0.25%	0.74%	0.50%

Source: Economic Insight review of 'Ofwat - top-down ODI rates - full models'.

We have also considered what effect including 2022-23 data has on Ofwat's above analysis. As shown in the table below, the mid-point for water PCs increases whilst the mid-point for wastewater PCs decreases. Whilst the results do not appear overly sensitive to the additional year of data, we note that a sample including only 2 or 3 years of data is relatively small – and would therefore objectively lack robustness.

[39] *'Top-down approach: Using customer preferences to inform ODI rates', Ofwat (June 2023).*

Table 3: Ofwat's corrected analysis to support its choice of average starting RoRE, compared to also including 2022-23 data

		Average UQ payment per PC	Average P90 payment per PC	Mid-point between UQ and P90
Ofwat's corrected analysis	**Water PCs**	0.35%	0.70%	0.53%
	Wastewater PCs	0.25%	0.74%	0.50%
Including 2022-23 data	**Water PCs**	0.42%	0.82%	0.62%
	Wastewater PCs	0.25%	0.69%	0.47%

Source: Economic Insight review of 'Ofwat - top-down ODI rates - full models'.

Ofwat's allocation of PCs to a starting RoRE

We have found that the alignment between the customer research Ofwat has relied on and the definitions of PCs is weak. The amount of RoRE Ofwat has initially allocated to each PC is therefore highly subjective.

Ofwat has relied upon three pieces of customer research to determine performance rankings for each PC. The surveys asked about aspects of service that differ from the precise definition of PCs, and therefore Ofwat has mapped survey questions to PCs. The table below illustrates how this mapping has worked for the Savanta customer research (labelled as 'Customer research 3' by Ofwat). As can be seen, the table below shows that 13 PCs are mapped to 5 separate survey labels. We find that the same survey label is mapped to significantly different PCs. For example, '*prevent sewage entering people's homes*' is mapped to both internal and external sewer flooding. In this scenario, the link to external sewer flooding is not clear.

Table 4: Savanta customer research response mapping

Savanta survey PC labels	PC that Ofwat maps to
Provide clean, safe drinking water	CRI, customer contacts
Prevent sewage entering people's homes	Internal sewer flooding, external sewer flooding
Prevent sewage entering the rivers, streams, and the sea	Serious pollution incidents, total pollution incidents, discharge permit compliance, storm overflows, river water quality, bathing water quality
Fix leaks	Leakage

Make sure there are no water shortages	PCC, business demand
Ensuring services can meet the needs of future generations	
Provide good customer service	
Keep bills low	
Reducing the amount of water taken from environmentally	
Reduce their carbon footprint	

Source: Economic Insight analysis of Savanta customer research.

By mapping one survey label to multiple PCs, it is also unclear whether customers value all PCs equally. Their response to the survey label could be influenced by a particular PC – and their valuation of the other PCs is significantly different. This is a greater issue for the labels that map to a bigger number of PCs.

Ofwat's choice of performance range for each PC

We have found that Ofwat's approach in relation to performance ranges lacks reliability and robustness. We also consider that greater clarity could be provided as to the purpose and expectation of the performance ranges – for example, whether Ofwat intends them to be reflective of all risks that actual companies face (e.g. including extreme weather events that happen infrequently, and poor performance).

The lack of reliability and robustness of Ofwat's approach is illustrated by the material changes in the incentive rates for certain PCs when the 2022-23 year of performance is included in Ofwat's approach (this data was not available when Ofwat conducted its analysis). More specifically, we have identified the following problems in relation to how Ofwat has calculated performance ranges:

- (i) unbalanced samples, whereby some companies contribute more observations and likely skew resulting performance ranges;
- (ii) small sample sizes, which means performance ranges are unlikely to be an accurate reflection of the full range of likely performance; and
- (iii) Ofwat's distribution for demand PCs (leakage, PCC, and business demand) is not based on sound logic and is inconsistent with Ofwat's wider approach.

We detail each of these issues further below.

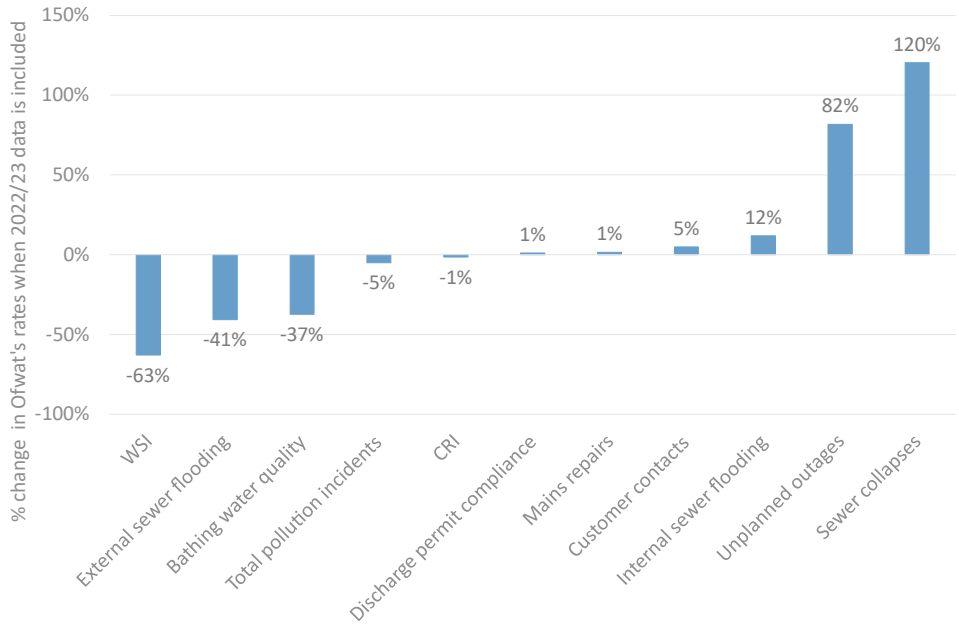
In addition, we also note that the choice of asset health performance range is arbitrary, because it is not based upon data. Rather than the P90 being informed by the distribution of historical performance, Ofwat selects a range using its 'judgement'. In

effect, this allows it to choose the ODI rate that it wants, and reverse-engineer this rate by choosing the performance range that generates this result.

Additional year of performance data

The figure below shows the change in incentive rates that result from including the 2022-23 performance data in Ofwat's top-down approach (where available).[40, 41] As can be seen, whilst some rates do not change significantly, rates for supply interruptions, external sewer flooding and internal sewer flooding change by more than 10%. Rates for asset health measures (mains repairs, unplanned outages, and sewer collapses) also change significantly, although we note that Ofwat's approach for these PCs is not as mechanical as the other rates.[42]

Figure 7: Change in Ofwat's indicative incentive rates when 2022-23 data is included



Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'.

The changes in incentive rates shown in the above chart are a product of the change in the P90 level of historical performance. These changes are shown in the table below for each PC.

[40] *2022-23 performance data has been taken from the industry datashare distributed in July 2023 and is subject to revision.*

[41] *We have collected the updated data from the APR 22-23, which only records data on PCs that were common at PR19. In addition, we have collected bathing water quality data from the Environment Agency website.*

[42] *For asset health measures, Ofwat sets the PCL based upon its 'reasonable performance' range. Rather than this performance range being simply informed by the data, Ofwat adjusts the range so that the incentive rate is at the level it wants. In our analysis in Figure 7, we assume that the absolute percentage adjustment Ofwat applied to each asset health PC would remain the same when new data is added. However, given Ofwat has adjusted the performance range to get targeted incentive rates, it is likely that Ofwat would simply apply different adjustments in light of the new data.*

Table 5: Summary of Ofwat's P90 ranges that are impacted by new data

PC	Is 2022-23 data available?	Percentage change in Ofwat's P90 range due to the inclusion of new data
Internal sewer flooding	Yes	-10.60%
External sewer flooding	Yes	68.96%
Bathing water quality	Yes	59.94%
Customer contacts	Yes	-4.70%
CRI	Yes	1.41%
Water supply interruptions	Yes	122.18%
Mains repair	Yes	-14.71%
Unplanned outage	Yes	-6.58%
Sewer collapse	Yes	-19.45%
Total pollution incidents	Yes	5.35%
Storm overflows	No	
Leakage	No	
PCC	No	
Business demand	No	
River water quality	No	
Discharge permit compliance – WaSCs	Yes	-1.11%
Discharge permit compliance - WoCs	No	
Serious pollution incidents	No	

Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'.

Unbalanced samples

The dataset that Ofwat uses to derive performance ranges does not include observations for all companies for all the included years. For example, for supply interruptions, Welsh only contributes two years, whereas some other companies contribute 6 years. This likely gives rise to a sample selection bias. That is, the results of Ofwat's analysis will likely be skewed because either high performers or low performers contribute fewer observations.

To investigate this, we assessed how Ofwat's P90 would change if it only used 'balanced' samples i.e. removing observations so that all companies contribute the same years to the sample. The table below shows which PCs have a balanced sample, and for those that do not, what the effect on the P90 estimate is of 'balancing' the sample. As can be seen, the P90 for supply interruptions changes significantly based on this balancing.

Table 6: Summary of which PCs Ofwat uses a 'balanced' sample of data to calculate

PC	Is the data 'balanced'	Does the P90 range change when balanced
Internal sewer flooding	No	Decreases from 95% to 86%
External sewer flooding	Yes	
Bathing water quality	No	Increases from 4% to 6%
Customer contacts	Yes	
CRI	Yes	
Water supply interruptions	No	Increases from 122% to 574%
Mains repair	Yes	
Unplanned outage	Yes	
Sewer collapse	Yes	
Total pollution incidents	No	Increases from 41% to 286%
Storm overflows	Yes	
Leakage	Yes	
PCC	Yes	
Business demand	Yes	

River water quality	Yes	
Discharge permit compliance – WaSCs	No	Decreases from 2.53 to 2.15
Discharge permit compliance - WoCs	Yes	
Serious pollution incidents	No	Decreases from 1.34 to 1.29

Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'.

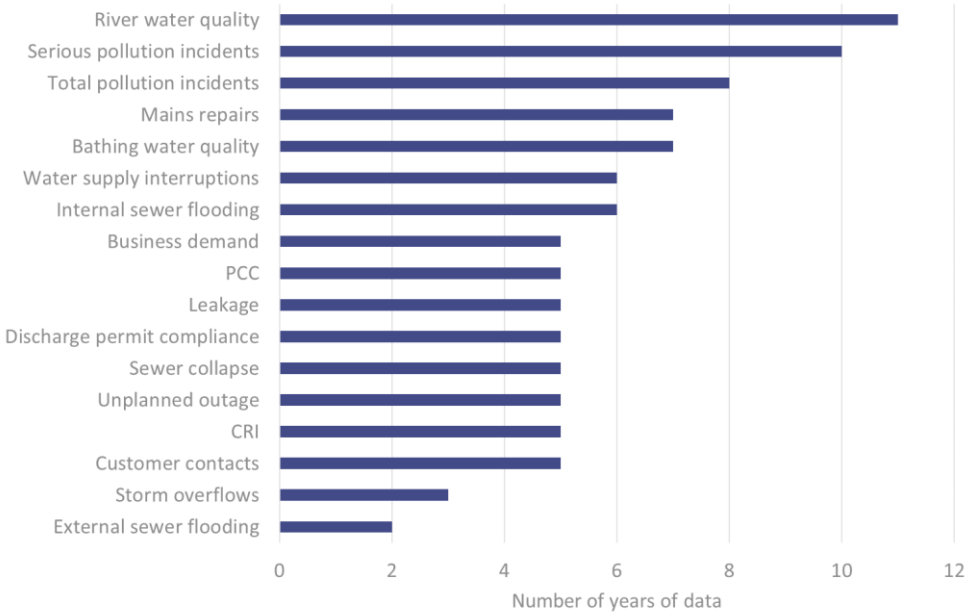
Such unbalanced samples result in a lack of robustness and reliability of Ofwat's approach. However, we recognise that although a balanced sample will not suffer from the identified sample selection bias, it will be a smaller sample – and therefore there is a trade-off between having a balanced sample and the number of observations.

Small sample sizes

Related to the above, Ofwat's approach also suffers from small sample sizes. Given the P90 is presumably intended to identify extreme levels of performance that occur infrequently, one cannot calculate an accurate estimate of it from small samples. For example, a small sample is unlikely to capture the effect of a 1-in-10 year weather event, if the sample only includes a few years of observations.

The figure below shows the number of years of observations Ofwat relies on for each PC. For example, it shows that the P90 for external sewer flooding is based on 2 years' worth of data (given 9 WASCs, this equates to 18 observations, as data for Thames and HDD is not recorded). Whilst it is difficult to say exactly how many years of data would be required to produce robust and reliable estimates, 2 years is unlikely to be sufficient.

Figure 8: Sample size used by Ofwat's approach to calculate performance ranges



Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'.

We recognise that Ofwat will have the 2022-23 year of data available to it for its draft determinations, and further still the 2023-24 year of data for its final determinations. However, consistent with criterion 4, continual changes to PR24 incentive rates will limit the ability of companies to make investment decisions based on them and damage regulatory certainty.

Distribution for demand PCs

The distributions for demand PCs (PCC, business demand, and leakage) are based on an aggregation. To aggregate, Ofwat sums the performance data of each of the three individual PCs. In the context of Ofwat's wider approach, we do not consider that this is an appropriate treatment for the demand PCs.

Ofwat's approach appears to hinge on the belief that there should be a single unit rate per Ml/d for PCC, business demand and leakage. For this to be appropriate, it would have to be that the marginal benefits and marginal costs (or the net of them) are the same for all three PCs. Whilst the wider social benefit of reducing each of these measures (reduced use of a scarce resource) may be similar at a 'national' level, the private cost (which is a component of social welfare in economics terms) of reducing these measures will likely vary. For example, the private cost of: reducing personal water consumption is the disbenefit of a long shower or washing the car; reducing business usage is the cost of the next best alternative in terms of a production input; and reducing leakage is the cost of fixing leaks / improved infrastructure. These all appear quite different.

There are also three inconsistencies in Ofwat's approach if it uses a single unit rate:

- Ofwat uses customer research to assign each PC an initial RoRE allocation, based on the 'importance' to customers. The resulting initial RoRE allocations are 0.4%, 0.5% and 0.6%, which Ofwat then sums in its calculation of the incentive rate. It is inconsistent to use varying customer importance ratings and then assume that the social benefit is the same.

- Ofwat has chosen to have separate PCs for each measure. If a standardised unit of performance of each measure is equivalent, then it would follow that there should be a single PC covering all three measures. Ofwat actively decided against this at PR24.[43]

- If Ofwat's general approach that it applies to other PCs was sound, and it was appropriate for there to be a single unit incentive rate for demand PCs, one would expect Ofwat's general approach to give broadly the same incentive rates compared to when all three demand PCs are aggregated. This is clearly not the case.

Furthermore, there are likely to be differences across companies in terms of their marginal benefits and marginal costs given: (i) their availability of water resources; and (ii) where they are on the cost curve.

We recognise that a single unit rate would incentivise companies to reduce aggregate demand through the most efficient channel – which is a positive feature of Ofwat's approach. However, as discussed above, this is inconsistent with setting separate PCs for each measure.

More specifically in relation to how Ofwat has calculated the incentive rates, we also disagree with the logic of aggregating historical performance and setting one performance range. Ofwat's logic for doing so is that covid increased PCC and reduced business demand. Whilst covid may have affected these measures, this is in essence what performance ranges do – capture the future uncertainties. It may be unlikely that another pandemic hits and has the same effect as covid, but with the ambitions to materially reduce consumption and England and Wales coming out of the pandemic era, high uncertainty and high performance volatility may well still remain.

Furthermore, Ofwat's logic of aggregating performance across all of PCC, business demand and leakage is flawed. Ofwat suggests that there has been an increase in household consumption and a reduction in non-household consumption driven by covid – but it does not suggest that leakage has been affected by covid. Therefore, on Ofwat's rationale (which we do not agree with), one would only aggregate across PCC and business demand.

Given the above, we have calculated the performance range for each PC using disaggregated data. We do this by calculating P10 / P90 values for each PC using its performance data. For example, for leakage, we calculate its P10 / P90 values using only leakage performance data.

[43] *'PR24 Final Methodology: Appendix 7 – Performance commitments', Ofwat (December 2023), page 51.*

The following table illustrates that disaggregating the performance range makes a significant difference to the size of the performance range. Notably:

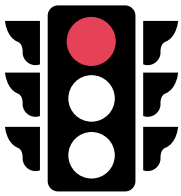
- The disaggregated performance range is twice as large for leakage, and approximately three times as large for PCC and business demand. This occurs because performance for these PCs move in opposite directions: PCC performance is skewed towards underperformance; whilst the data for the other two PCs is skewed towards outperformance. Aggregating the performance range means these 'skews' cancel each other out – giving the impression there is less movement in the data than is actually the case for each individual PC.
- Disaggregating changes whether the underperformance or overperformance side of the distribution is selected for two PCs. Whilst the underperformance data was used for aggregate demand, we find that two PCs (leakage and business demand) use overperformance data when disaggregated.

Table 7: Illustration of how the performance data differs for the demand PCs when disaggregated demand is used instead of aggregated demand

PC	Aggregated	Disaggregated
Leakage	-5%	10%
PCC	-5%	-14%
Business demand	-5%	16%

Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'. Note: numbers denoted as negative illustrate that the underperformance side of the distribution has been used to select the performance range, whilst positive numbers indicate that the overperformance side of the distribution has been used.

Conclusion



Overall, Ofwat's approach does not meet this criterion. Its approach does not calculate either a reliable RoRE at risk or performance range estimates. In summary, the key flaws with Ofwat's approach are that it:

- (i) relies upon arbitrary assumptions and flawed logic for key elements of its approach; and
- (iii) relies upon data samples that are unbalanced and too small, meaning that its results are not reflective of the performance risk that companies actually face.

We therefore rate Ofwat's approach red in relation to this criterion.

We also recognise that Ofwat is restricted by the data that it has access to. However, we caution against Ofwat updating its approach each time new data is available. As per criterion 4, it is detrimental to continually change incentive rates by a large extent.

3D. Criterion 4: The approach does not produce significantly different incentive rates from PR19

Ofwat's approach

Ofwat's approach generally does not directly take account of the incentive rates from PR19, nor has Ofwat presented an analysis of how its indicative PR24 rates compare to those from PR19. The PR19 rates do feature in terms of Ofwat's PR24 approach in terms of the following.

- The PR19 rates determine the hypothetical incentive payments that Ofwat has used in its analysis to support its choice of 0.5% RoRE at risk for PCs on average. As shown subsequently in this section, this does not result in Ofwat's indicative PR24 rates aligning with those it set at PR19.
- Ofwat targeted higher incentive rates for mains repairs and sewer collapses compared to the PR19 rates. It did this through adjusting the performance range it used to spread the RoRE over.

Our assessment

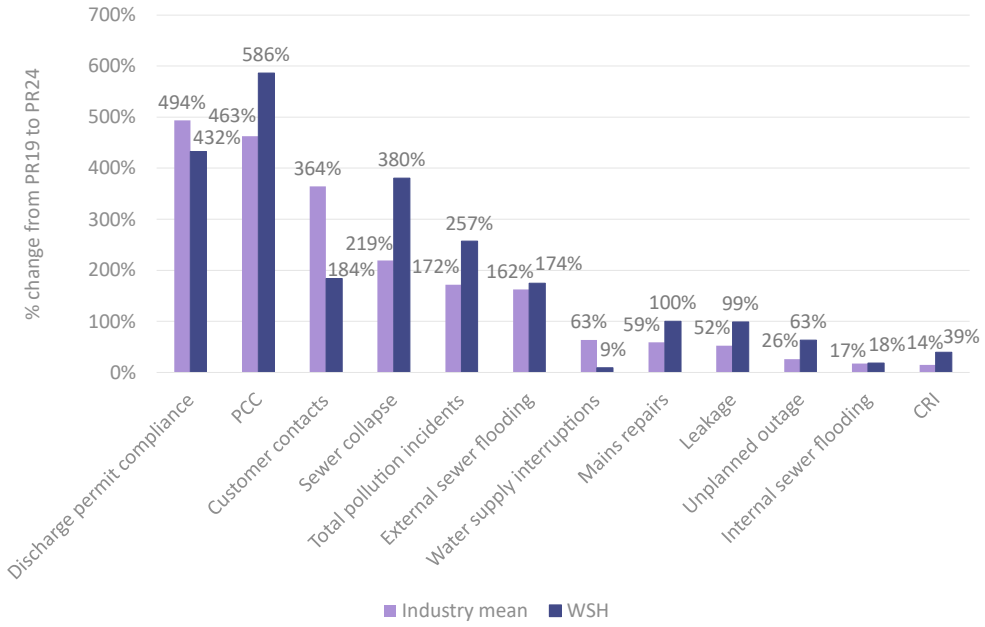
Ofwat's indicative PR24 rates are materially different to those it set at PR19, and therefore Ofwat's approach does not meet this criterion.

Figure 9 below shows the difference between PR19 rates and Ofwat's indicative PR24 rates, for Welsh specifically and then for the average industry rates. The PR19 rates have been adjusted for the price base and normalisation parameters.[44], [45] As can be seen, all of the indicative PR24 rates are higher than at PR19 – with many of the rates being more than 100% larger.

[44] *For example, the internal sewer flooding ODI rate is reported as the number of internal sewer flooding incidents normalised per 10,000 sewer connections. To apply normalisation parameters, we calculate a unit rate (i.e. the rate per sewer flooding incident) by dividing the PR19 ODI rate by the number of sewer connections used to calculate PR19 rates. To make this comparable to PR24 rates, we estimate a normalised PR24 comparable ODI rate by multiplying the unit rate by number of sewer connections Ofwat has used to calculate its PR24 rates.*

[45] *Due to data limitations, the PR19 ODI rates for discharge permit compliance and unplanned outage have not been adjusted for changes in normalisation parameters between PR19 and PR24. We do not expect they have changed significantly between price controls.*

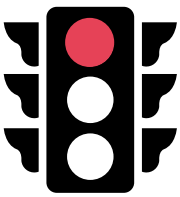
Figure 9: Percentage change from PR19 rates to Ofwat's indicative PR24 rates – Welsh Water and industry-wide



Source: EI analysis of Ofwat's indicative PR24 rates against equivalent PR19. Note, PR19 rates have been adjusted for the price base and normalisation parameters.

As discussed in section 2E, we do not consider there are any legitimate reasons for incentive rates to increase significantly compared to PR19. Furthermore, Ofwat has not provided any explanation as to why rates its rates are significantly higher.

Conclusion



We consider that Ofwat's approach does not meet this criterion. We find that there is a significant difference between the PR19 and PR24 incentive rates. We therefore rate Ofwat's approach red in relation to this criterion.

3E. Criterion 5: The approach is consistent with a balanced package of risk

Ofwat's approach

Ofwat does not provide a view on whether its indicative rates contribute to a balanced package of risk. The elements of Ofwat's approach that are relevant to the balance of risk are as follows.

- **Ofwat calculates an 'additive' amount of RoRE at risk**. Specifically, Ofwat adds up the starting amount of RoRE for each PC, which gives 4-5% RoRE for water and 3.5-4.5% RoRE for wastewater.[46] It states that this is not the level of risk that a company faces because there will be a mix of out- and under-performance,[47] with the target level of RoRE for the outcomes package being ±1% to ±3% for an efficient firm.[48] Ofwat does not perform any further risk analysis.

- **Ofwat sets symmetrical incentive rates using either the P10 or P90 of historical performance for each ODI.** Specifically, Ofwat selects either the historical P10 or P90 performance range from the PCL for each ODI, according to which is larger in absolute terms, before splitting the ±0.4% to ±0.6% RoRE (depending on the perceived customer importance of the ODI) over this range to get both out- and under-performance rates.

- **Ofwat limits the use of bespoke ODIs at PR24**, concentrating instead on common ODIs across the industry. It is expected that there will be fewer bespoke PCs at PR19 compared to PR24.

Our assessment

We find that Ofwat's approach likely contributes to a **negatively skewed** package of risk. We also find that Ofwat's approach likely gives rise to **greater risk** from the ODI package, compared to PR19. We detail our findings below.

Negatively skewed package of risk

As discussed in section 2F, the balance of risk of an outcomes package will depend on both the incentive rates and expected performance relative to where the PCLs are set. There is also a distinction between actual companies and the notional company. We find that Ofwat's approach likely contributes to a negatively skewed package of risk for both actual and notional companies, for the reasons set out below.

First, Ofwat's indicative incentive rates are symmetrical but the expected level of performance for a notionally efficient company is likely to be below where Ofwat will set the PCL. Whilst Ofwat has not yet provided a view on what PCLs it will set at PR24, its final methodology (and PR19 approach) are based on extrapolating industry

[46] *'Top-down approach: Using customer preferences to inform ODI rates', Ofwat (2023), page 3. The ranges reflect the fact that the number of PCs will depend on whether a company is a WaSC or WoC.*
[47] *'Top-down approach: Using customer preferences to inform ODI rates', Ofwat (2023), page 3.*
[48] *'PR24 Final Methodology: Appendix 10 – Aligning risk and return', Ofwat (December 2023), page 15.*

performance for each individual PC – rather than what an efficient firm can actually deliver. Ofwat's approach fails to recognise the relationship between PCs (an efficient firm cannot 'outperform' on all PCs), diminishing marginal gains / increasing marginal costs, and exogenous differences between companies.

Second, abstracting from where exactly the PCL is set, the performance distribution for an efficient company will likely be negatively skewed. This is for two reasons.

- It is more likely that an efficient company will 'underperform', for example due to extreme weather, than it will 'outperform'. This will increasingly be the case as the level of performance of the efficient firm increases, because the scope for outperformance will become more limited.

- There are natural limits to outperformance (e.g. zero supply interruptions) and, relatedly, there are two penalty-only PCs where it is not possible to outperform at all (CRI and discharge permit compliance).

To quantify the latter point, we have assessed how much of the RoRE at risk from Ofwat's approach is not achievable. More specifically:

- We have identified that for three PCs, Ofwat's P90 level of performance is not achievable. For example, for supply interruptions, Ofwat uses a P90 level of performance of 122%. Based on the 2024-25 PCL, this suggests a performance level of -1.08 minutes, which is clearly not a possible level of performance. This results in 0.11% of the potential 0.6% RoRE upside not being possible (for the median firm).[49] Table 8 below shows the amount of RoRE that is not possible for each of the three identified PCs.

- For the two penalty-only PCs, CRI and discharge permit compliance, we take the full 0.6% and 0.5% RoRE at risk.

- As shown in Table 8, in total 0.82% RoRE is not available for water PCs, and 0.63% RoRE is not available for wastewater PCs.

[49] *For each individual PC, the median firm is defined as the firm whose unit rate is applied to all other firms. Ofwat calculates this unit rate by dividing the equity at risk by the performance range for each firm. The median firm is the firm whose unit rate represents the median value of all the companies' unit rates. This means that the median firm will have Ofwat's target percentage RoRE at risk for that PC, whilst all other firms will have different percentages of RoRE at risk.*

Table 8: RoRE feasibly at risk for PCs with an unachievable P90 range

PC	RoRE not available
Water supply interruptions	0.11%
Customer contacts	0.11%
CRI	0.60%
Total water RoRE unavailable	**0.82%**
Serious pollution incidents	0.13%
Permit compliance	0.50%
Total wastewater RoRE unavailable	**0.63%**

Source: EI analysis of 'Ofwat - top-down ODI rates - full models'.

Using an additive approach to risk analysis and Ofwat's performance ranges, the total potential downside from Ofwat's indicative rates is 4-5% RoRE for water PCs and 3.5-4.5% RoRE for wastewater PC (for a hypothetical median firm).[50, 51] However, as set out above, the upside is curtailed by natural limits. Therefore, this suggests that the potential upside is only 3.18-4.18% RoRE for water PCs and 2.87-3.87% for wastewater PCs. These results are presented in the table below, which demonstrate that Ofwat's indicative incentive rates will contribute to a negative skew in the ODI package for the efficient firm.

Table 9: Available upside and downside based on 'additive' approach and Ofwat's performance ranges, for a hypothetical median firm

	Water PCs	Wastewater PCs
Upside	3.18-4.18%	2.87-3.87%
Downside	4-5%	3.5-4.5%

Source: Economic Insight analysis

Third, historical ODI payments and company performance suggest that Ofwat's approach to PCLs and ODIs will give rise to a negatively skewed package of risk for actual companies.

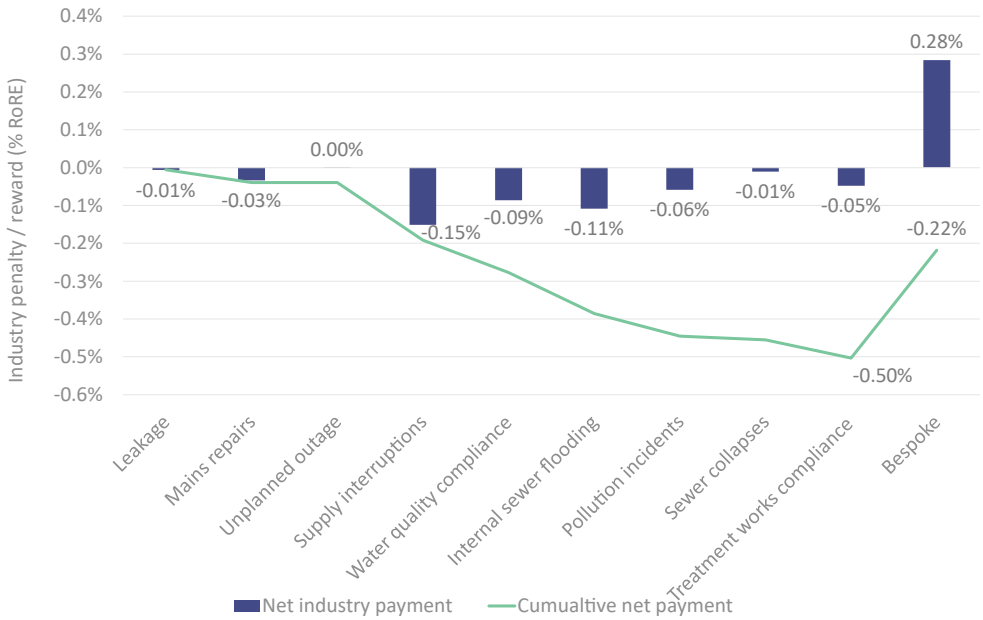
Figure 10 below shows the industry ODI payments over the first three years of AMP7. As can be seen, common PCs have typically attracted negative ODI payments (and we note that the industry position on bespoke ODI payments is largely driven by two

[50] *This is the same as Ofwat's view, as set out in 'Top-down approach: Using customer preferences to inform ODI rates', Ofwat (June 2023).*

[51] *The hypothetical median firm is a firm that is the median firm for every PC i.e. it will have Ofwat's target percentage RoRE at risk for every PC.*

companies). Given the ODI package at PR24 will predominantly consist of common PCs, this suggest that there will be a negative RoRE skew for actual companies at PR24.

Figure 10: Industry-wide RoRE at risk for ODIs in the first three years of AMP7 for common and bespoke PCs

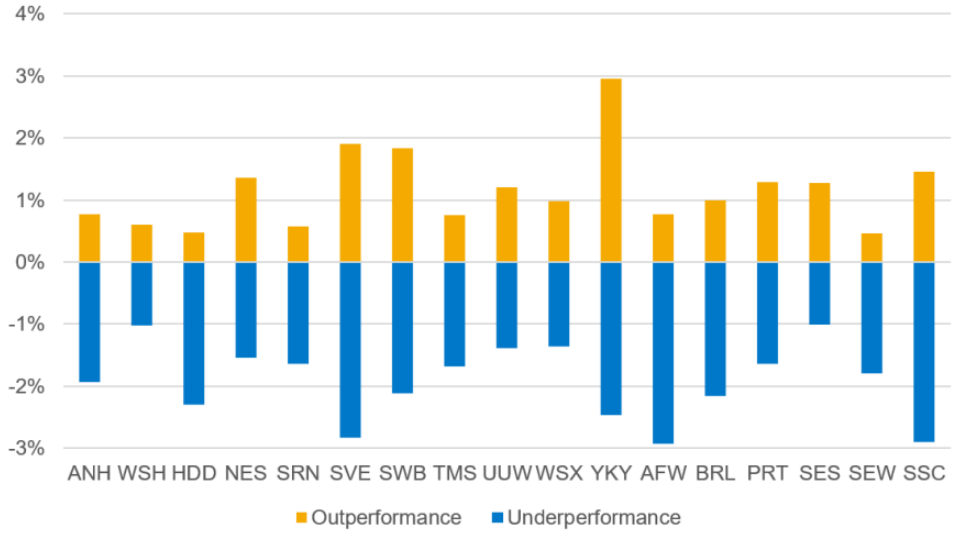


Source: Economic Insight analysis of Ofwat's WCPR and APR data. Note: PCC is not included in this analysis because it is to be determined at the end of PR19; % RoRE figures are annual averages.

Overall RoRE at risk

Our starting point for considering the totality of RoRE at risk from Ofwat's ODI approach is Ofwat's view of ODI RoRE risk at PR19. As illustrated below, Ofwat's view was that ODI risk across the industry was consistent with its view of ±1% to ±3% RoRE. Some companies are towards the lower end of the range, whereas others are towards the upper end. Based on this, and that Ofwat's view that the package should still be around ±1% to ±3% RoRE, we would not expect to see significant movements in RoRE at risk.

Figure 11: Ofwat's PR19 ODI RoRE risk ranges



Source: 'PR19 final determinations: Aligning risk and return technical appendix', Ofwat (December 2019), page 28, Figure 3.6

To assess how RoRE at risk has changed, we consider three analyses, as detailed below. We also assess how the amount of risk can be assessed by converting 'additive' risk into 'real' risk.

Intuitive consideration

All else equal, increasing the incentive rates will increase the RoRE at risk. As shown in section 3D, Ofwat's indicative incentive rates are all higher than the average of companies' incentive rates from PR19 (adjusted to ensure like-for-like comparison). Furthermore, in a material number of instances the rates are greater than 100% larger (i.e. more than double), and in some instances materially higher still.

Compounding the increase in incentive rates, we note that there will be more common PCs at PR24 compared to PR19. Including more PCs with relatively high incentive rates will further increase RoRE at risk.

However, counteracting the above, the number of bespoke PCs will be reduced at PR24. Furthermore, we recognise that: (i) there is likely to be real RCV growth over the PR24 period, thus increasing regulatory equity; and (ii) Ofwat has decreased the notional gearing ratio from 60% to 55%, thus also increasing regulatory equity (although such a change in regulatory assumption does not instantaneously change anything in the 'real world'). Increasing regulatory equity will, all else equal, decrease RoRE (because regulatory equity is the denominator in the calculation of RoRE). To illustrate the effect of these changes, a hypothetical 10% increase in RCV and 5 percentage point decrease in gearing would result in a % RoRE decreasing by roughly a fifth (20%).[52]

[52] *For example, assume a return of £5, an RCV of £500 and gearing of 60%. This would result in a RoRE of 2.5% (5 / (500*(1-60%)). Then assume the RCV increases by 10% and gearing reduces to 55%. This would result in a RoRE of about 2.0% (5 / (550*(1-55%)). Hence, % RoRE would reduce by about a fifth (20%).*

Whilst it is difficult to draw firm conclusions from this, given the magnitude of the increase in incentive rates, it suggests to us that the total amount of RoRE at risk has increased between price controls.

PR19 performance based on PR19 and PR24 incentive rates

Secondly, we have compared: (i) actual PR19 ODI payments; with (ii) the equivalent ODI payments based on Ofwat's PR24 indicative rates. We have calculated the latter by:

- for each company for each common PC, calculating the percentage change between the PR19 rate and Ofwat's indicative PR24 rate;[53]
- applying the above percentage changes to the PR19 ODI payments for each company for each PC for each year;
- summing the above monetary payments for each PC for each year, and then dividing by the sum of industry regulated equity for that year; and
- taking the average across the three years.[54]

The following table shows the results. As can be seen, applying PR24 rates to PR19 performance data increases the common PC payments the industry would have accumulated, from -0.50% RoRE to -0.70% RoRE per year. Furthermore, whilst common PC payments have been negative over the first three years of the PR19 period, they have been offset by positive bespoke PC payments (albeit this is largely driven by two companies). Without bespoke PCs (as will broadly be the case at PR24), there would be nothing to counteract the negative common PC performance. This analysis therefore suggests that Ofwat's indicative incentive rates will increase risk.

Table 10: Comparison of ODI payments when PR19 and PR24 rates are applied to PR19 outturn data – industry-wide, % RoRE

	PR19	PR24
Common PC payments	-0.50%	-0.70%
Bespoke PC payments	0.28%	
Total payment	-0.22%	-0.70%

Source: Economic Insight analysis of WCPR data

We recognise that the above estimates do not take into account any increase in regulatory equity over the PR24 period, either because of lower notional gearing or RCV

[53] *This percentage change is calculated on a like-for-like basis in terms of the price base and normalisation factors (e.g. number of connections).*

[54] *We have note that this method scales PR19 payments, rather than calculating payments 'bottom-up'. As a result, we do not take account of the levels at which caps and collars were set at PR19.*

growth. However, regulated equity would need to change by much more than expected to change the conclusion from this analysis that risk has increased.

Additive risk ranges

Thirdly, we have compared additive risk ranges for companies between PR19 and PR24. To do so, we have:

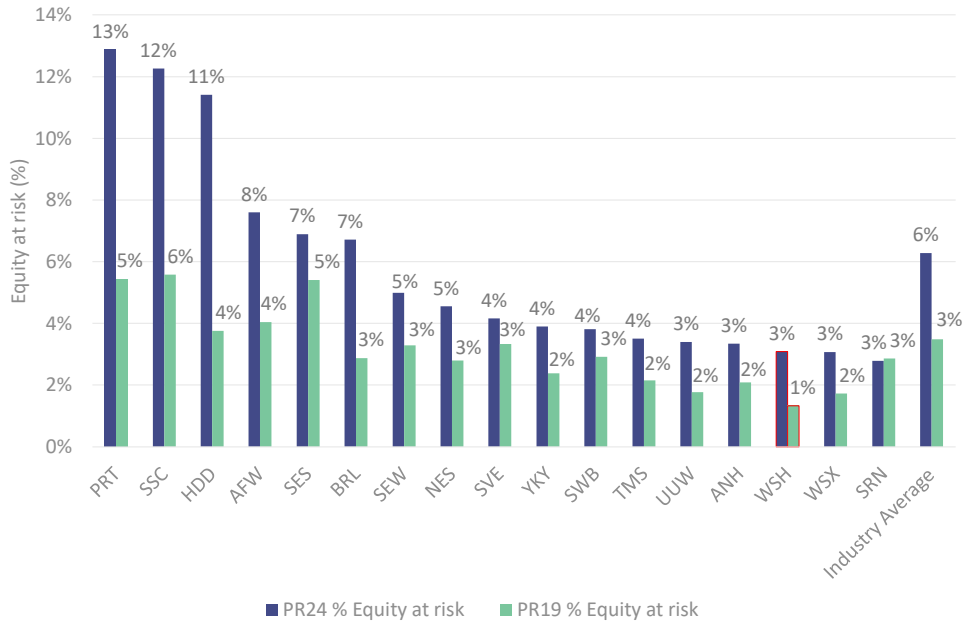
- adjusted PR19 ODI rates to be comparable to PR24 rates in terms of price base and normalisation parameters (e.g. number of connections);
- calculated the monetary value of ODI payments that companies would be liable for if they performed at Ofwat's P90 levels for performance, for each PR24 common PC where there is a PR19 equivalent; and
- divided this monetary value by total regulatory equity to convert figures into a percentage of RoRE.[55]

This analysis will be an underestimation of additive risk, because only PCs that are present in both price controls are included.

The following figure illustrates that the additive RoRE at risk has increased markedly between PR19 and PR24, for the PCs present in both price controls. For Welsh, the RoRE at risk has increased from 1.3% to 3.1%. For the industry in general, the additive risk has approximately doubled. We note that this analysis does not take any changes in regulatory equity into account. Nevertheless, as above, regulated equity would need to change by much more than expected to change the conclusion from this analysis that risk has increased.

[55] *We used 2022-23 regulatory equity figures, consistent with Ofwat's calculation of indicative incentive rates.*

Figure 12: Additive RoRE at risk over Ofwat's industry-wide P90 range for PR19 and PR24, for PCs present in both price reviews



Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'. Note: PR19 incentive rates have been adjusted to be comparable to PR24 rates in terms of price base and normalisation parameters (e.g. number of connections).

'Additive' risk to 'real' risk

In addition to the above, we have assessed what the additive amount of RoRE at risk means for the 'real' amount of RoRE at risk. This takes account of the fact that extreme levels of performance are unlikely to occur simultaneously across all PCs.

At PR19, Ofwat used scaling factors to convert the additive risk of ODI packages to 'real' risk. Ofwat applied a 90% scaling factor to outperformance payments and a 70% scaling factor to underperformance payments.[56] Whilst we do not endorse the validity of Ofwat's scaling factors, they can at least be used to provide a view of real risk based on Ofwat's own analysis.

Similar to the above analysis, to calculate Welsh's additive risk, we calculate monetary payments for each PC at the P90 level of performance and then divide by Welsh's regulatory equity.[57] We then sum the RoRE at risk across all PCs. This results in 4.65% RoRE at risk for Welsh.

Applying the scaling factors, as shown in the table below, we find that Ofwat's ODI rates result in Welsh facing a 'real' risk that is outside Ofwat's ±1% to ±3% target range for the notionally efficient company.

[56] *'PR19 final determinations: Delivering outcomes for customers policy appendix', Ofwat (December 2019), page 175.*

[57] *We used 2022-23 regulatory equity figures, consistent with Ofwat's calculation of indicative incentive rates.*

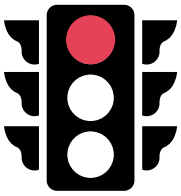
Table 11: Estimate of the 'real' risk of Ofwat's indicative ODI rates

	RoRE at risk
Additive risk for all Welsh PCs	4.65%
90% scaling factor	4.18%
70% scaling factor	3.25%

Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'.

If there is material RCV growth over the PR24 period, this will reduce the % RoRE at risk. However, the above analysis suggests that Welsh would still be around the top of Ofwat's risk range.

Conclusion



Overall, Ofwat's approach does not meet this criterion. Its approach likely contributes to a negatively skewed package of risk. The factors in Ofwat's approach that have driven this are:

- (i) Ofwat's incentive rates are symmetrical but the expected level of performance for a notionally efficient company is likely to be below where Ofwat will set the PCL;
- (ii) the performance distribution of an efficient company is likely to be negatively skewed; and
- (iii) historical data suggests that company payments are usually negative for common PCs.

We also find that the totality of risk is likely to be higher than it was at PR19. This is a result of materially higher incentive rates.

Given the above, we rate Ofwat's approach red in relation to this criterion.

We also note that the risk associated with Ofwat's indicative incentive rates (or any other set of incentive rates) could be further assessed through Monte Carlo analysis. This would allow for a direct estimate of risk ranges that does not rely on an 'additive' approach to risk or scaling factors. Ofwat has not provided such analysis, and it is beyond the scope of this report.

4 Alternative approaches

Following our detailed evaluation of Ofwat's current proposed indicative incentive rates (using its top-down approach) in Chapter 3, in this chapter we provide an assessment of the alternatives. In turn we:

- set out an overview of the alternative options identified and considered, within the scope of this work;
- provide our evaluation of the alternatives, using the same framework as detailed in Chapter 2 of this report; and then
- make recommendations for Welsh, in light of our findings.

4A. Overview of alternatives

In practice, a wide range of approaches could be used to set incentive rates. However, as the scope of our work is to advise Welsh as to what rates it could consider proposing within its PR24 Business Plan, we have intentionally limited the range of options to those that are practically feasible for Welsh at this point in the process. We note that, for other companies, the practical alternative options may vary.[58]

We therefore consider there to be four main options:

- **Ofwat bottom-up.** Apply the incentive rates as per Ofwat's collaborative customer research using a 'bottom-up' approach.
- **Ofwat top-down.** Apply Ofwat's current proposed indicative incentive rates (i.e. using Ofwat's top-down approach, as evaluated in the previous chapter).
- **Amended top-down.** Apply an amended 'top-down' approach, designed to mitigate some of the limitations with Ofwat's proposals as identified in Chapter 3.
- **PR19 rates.** This could be keeping Welsh's own incentive rates from PR19, or adopting an average across companies. Some adjustments would be applied, such as updating the price base and reflecting growth in e.g. number of connections.

We also consider that a hybrid of, for example, an amended top-down approach and the PR19 rates could be used. Nevertheless we consider the options in isolation for assessment purposes.

[58] *For example, Welsh has not undertaken its own new customer research for PR24, as it proceeded on the basis that Ofwat's collaborative customer research would be utilised under a bottom-up approach; whereas we understand some companies did commission new research – as such the option to use 'new' company specific research is not available to Welsh.*

4B. Evaluation

In this section, we set out our evaluation of the alternative options, relative to Ofwat's current proposed top-down method. The following table provides a 'traffic light' (RAG) rating of the options against our evaluation framework, along with the key reasons for our assessment.

Table 12: Summary of our evaluation and key reasons

Option / rating	Criterion 1:likely to create appropriate incentives		Criterion 2: takes into accounts whether there are material differences in costs and benefits across companies		Criterion 3: Estimation method is robust and reliable		Criterion 4: does not produce significantly different incentive rates from PR19		Criterion 5: produces symmetrical balance of risk (for efficient firm)	
	RAG rating	Key reasons and evidence / analysis	RAG rating	Key reasons and evidence / analysis	RAG rating	Key reasons and evidence / analysis	RAG rating	Key reasons and evidence / analysis	RAG rating	Key reasons and evidence / analysis
Ofwat bottom up		• Based on valuation of benefits • Does not reflect marginal costs		• Approach is capable of reflecting differences • No differences reportedly found		• Reliability and mapping issues		• Generally produces significantly larger rates		• Likely contributes to asymmetric package
Ofwat top-down		• Not designed to create appropriate incentives		• No variation between companies		• Various issues identified with sample and treatment of certain PCs		• Generally produces significantly larger rates		• Likely contributes to asymmetric package
Amended top down		• Not designed to create appropriate incentives • Can be tailored to specific targets		• No variation between companies, although could be added		• Some issues from Ofwat's top-down approach can be addressed		• Rates can be broadly aligned with PR19		• Could allow for symmetry issue to be addressed
PR19 incentive rates		• PR19 rates were generally based on marginal benefit and cost values		• In theory reflects differences, but depends on if company-specific or average rates are taken • Questions raised about robustness of differences		• Questions raised about robustness of rates		• Consistency by design		• Rates are asymmetric, but PR19 may not have been balanced

4C. Recommendations for Welsh Water

In light of the evidence and analysis contained in this report, and our independent assessment of the practical options available to the company at this time, our recommendations to Welsh are as follows.

- Firstly, we do not think Welsh should adopt Ofwat's indicative ODI rates *as currently proposed* under the regulator's top down methodology without revision. This is primarily because:
  - the implied incentive rates are not likely to create appropriate incentives and so may encourage Welsh to allocate its resource, time, effort and money in ways that do not best serve the interests of its customers, society, or the environment;
  - the method used by Ofwat under its top-down approach is currently insufficiently robust and reliable;
  - Ofwat's proposed incentive rates for Welsh are materially different to those that applied at PR19, which will further increase investor risk and costs to Welsh's customers in the long-run.
- Secondly, and notwithstanding the above, we have identified ways in which Ofwat's top-down approach could be amended in the near-term and would encourage Welsh to provide constructive feedback to the regulator, consistent with Ofwat's request in its letter to companies of June 2nd. Welsh should encourage Ofwat to reconsider its approach in light of this feedback.
- Thirdly, given the available options at this time, within its PR24 Plan, Welsh could propose incentive rates that are either: (i) based on those from PR19; or (ii) calculated from an amended top-down approach.
  - The advantage of using the PR19 incentive rates is, being primarily derived bottom-up, they are (in principle) consistent with creating incentives for the company to prioritise its efforts in a way that will benefit its customers. They also avoid any material increase in investor risk that could further increase costs to customers over time. Welsh could adopt either its own PR19 incentive rates or the industry average – the latter having the potential for consistency across the industry.[59] Adjustments would need to be made to the PR19 incentive rates to account for inflation and growth in, for example, customer numbers / connections. An alternative approach would also be required for PCs that are new at PR24.

[59] *We note that at PR19, Welsh generally adopted a top-down approach in the incentive rates that it proposed. These were subsequently subjected to Ofwat's triangulation approach / application of a 'reasonable range'.*

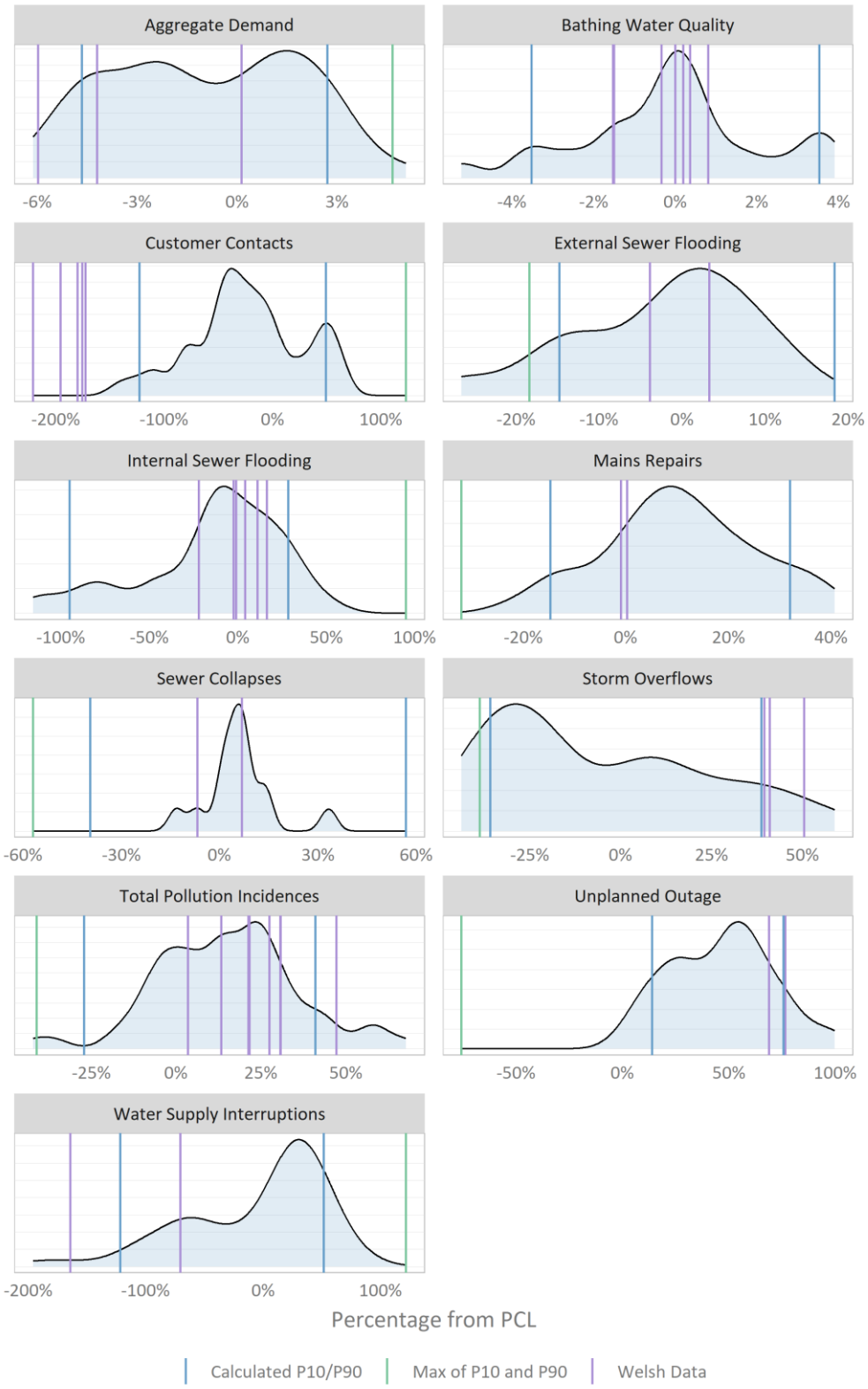
  - The advantage of the amended top-down approach is that it achieves consistency across companies, whilst can still result in incentive rates for Welsh that are relatively consistent with those used at PR19. It can be used to target a total amount of RoRE at risk, and it can also provide for a greater degree of judgement in setting rates – although this should be exercised with caution. It can be used for both new and existing PCs.
  - It could further be appropriate for Welsh to triangulate across the above two options, either 'across the board', or more selectively by PC (i.e. choosing whichever it felt most suitable, in light of our criteria, PC-by-PC).

- Fourthly, Welsh should consider, and highlight, key interdependencies across its Plan. Specifically, noting that, should Ofwat not revise its approach to incentive rates and 'impose' its view in its PR24 determinations, this may affect Welsh's view on the appropriate cost of equity (i.e. if the incentive rates give rise to an expectation of a downside skew in equity returns).

5 Annex

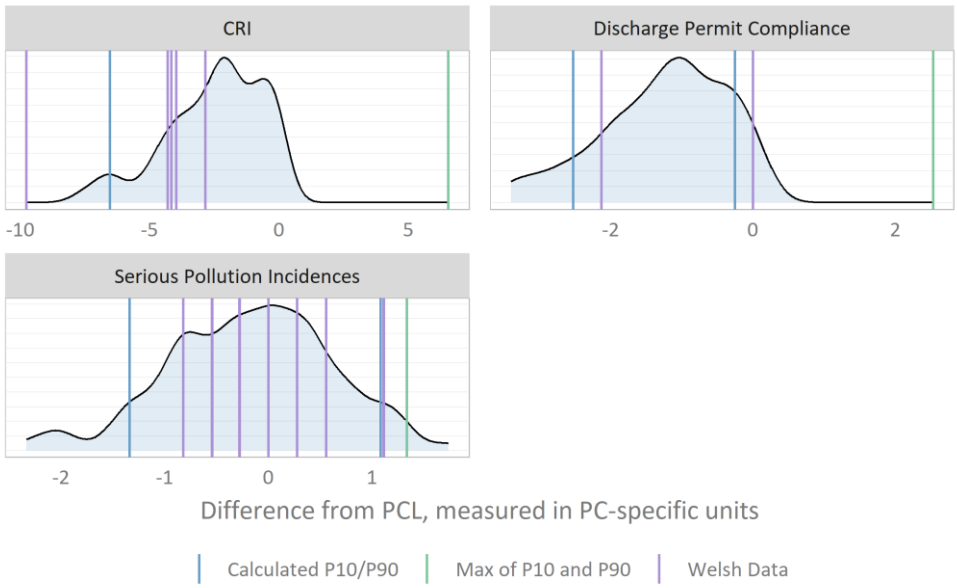
The following two figures illustrate the performance distributions for each PC that Ofwat used in its calculations of the indicative incentive rates. The blue lines show the P90 and P10; the green lines show the maximum of the P90 and P10 imposed on the other side of the distribution; and the purple lines show observations for Welsh. PCs for which the performance distributions are measured in percentage deviations from the PCL are shown separately to the those measures in PC-specific units.

Figure 13: Illustration of Ofwat's performance range distributions, for PCs for which distributions are measured as 'Percentage from PCL'



Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'. Note: for bathing water quality, the P10 and P90 are the same in absolute terms, so there is no green line.

Figure 14: Illustration of Ofwat's performance range distributions, for PCs for which distributions are measured in PC-specific units



Source: Economic Insight analysis of 'Ofwat - top-down ODI rates - full models'.

WE MAKE ECONOMICS RELEVANT

Economic Insight Limited

125 Old Broad Street
London
EC2N 1AR
+44 207 100 3746
www.economic-insight.com