

Outcomes Working Group

Peter Jordan, Jeevan Jones

13 January 2022



Agenda

Aim: to discuss options to determine ODI rates where we do not have direct customer evidence, such as for asset health PCs

- 10:00 Welcome
- 10:05 Presentations
- 10.40 Breakout session
- 11.10 Feedback
- 11.30 Conclusions and look ahead
- 11:40 Close



Operational resilience and asset health PCs

November Performance commitments for future price reviews discussion paper

December Assessing base costs at PR24 discussion paper (includes asset health measures)

January **This discussion**
Responses to discussion papers (by 3 February for base costs)

February Final report from UKWIR Future Assets project

March Further discussion on asset health and operational PCs

Setting ODI rates: three approaches...

Marginal benefit-based

Potentially with a simplified ODI rate formula: $MB^*50\%$

Use willingness to pay (WTP) or other valuation methods

Cost-based

Where we lack strong evidence of willingness to pay?

Reliability of data?

Top-down allocations

Allocate maximum payments for individual or groups of PCs





BRISTOL
WATER



**How could we
simplify ODI rate
setting?**

bristolwater.co.uk

1. At PR14 & PR19, incentive rates and service levels were set “bottom up”, using a market proxy where marginal cost = marginal WTP for the improvement.

Even though stated preference customer research surveys were used to obtain some values used in plan development, and to inform the incentive rates set based on marginal cost = marginal WTP (adjusted for totex cost sharing incentives between customers and companies), a single survey result would rarely be applied without a wider context. Triangulation described the general principle of considering a range of incentives. For instance, Bristol Water used a scenarios approach to triangulate service levels for different bill contexts through customer research, using this principle.

2. This approach to incentives reflected local company specific service levels, with trade offs through customer research and engagement used to inform both company plans and ODI design.

Differences in service levels and priorities meant RoRE ODI rates could vary. In any case the risk and return impact varied by company, including where Ofwat compared and partly standardised incentive rates, design and service levels at PR19.

3. With standardised interventions on PC levels and ODI rates and design at PR19, Ofwat sometimes had to constrain its interventions in order to maintain a risk and return balance,

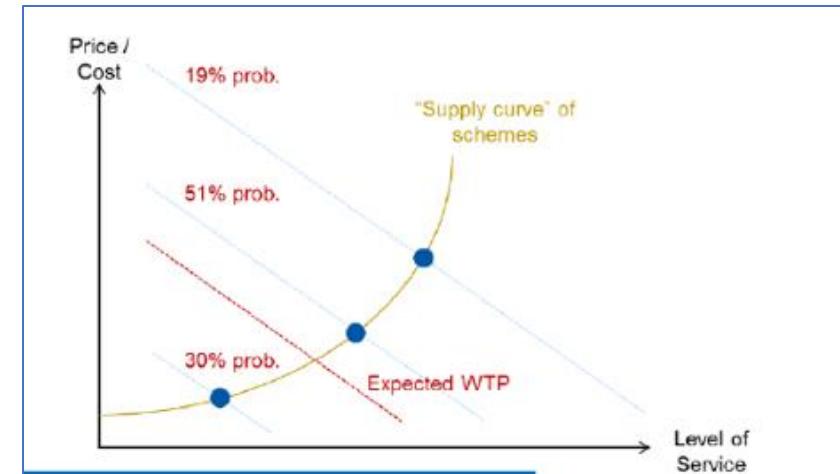


Figure 3.6: PR19 ODI risk ranges calculated as a percentage of regulatory equity



Source: Ofwat

4. For PR24, the initial design suggests more standardised packages of outcome incentives. Ofwat intend to specify the service level consistent with base expenditure. Variations in service levels for individual companies could result in cost adjustments to base expenditure (which could therefore be either positive or negative). Changes in future service levels from a standard level would be considered as enhancement expenditure. For enhancement expenditure and service levels, Ofwat may set output related Price Control Deliverables would be used, which may reduce some use of local bespoke ODIs and incentive variation between companies that featured at PR19.

The PR14 & PR19 approach to setting incentives based on marginal benefit less marginal cost (adjusted using a totex sharing rate) may therefore not apply. Customer research would still be required to set the ODI rates for the outcomes that are consistent with base expenditure allowances Ofwat set, and a cost sharing rate can be assumed to take into account costs (e.g. the incentive would be 50% of the benefit value, if the totex sharing rate was 50%). Based on the evolving list of potential performance commitments, many of the metrics, particularly on asset health, have not featured directly in past company WTP stated preference customer research.

5. For risk & return, there are a number of reasons why the incentives package may need to vary between companies. It is possible that the long-term environmental and water resource destination for companies may vary. The pressures on the water service may be different from wastewater. Companies have different historical levels of service and network resilience, in particular for key aspects such as leakage and mains bursts. Innovation and other local factors may mean that for some companies, an efficient plan for their customers may suggest a different direction. It is a hard task for regulators to reflect all of these factors without the burden (and risk from information asymmetry) in scrutinising company plans in a fully bespoke fashion. It is however difficult to design valid customer research without exploring where such local differences matter to customers and stakeholders. Profiles of efficient expenditure and delivery can vary by company over time. To reveal information, Bristol Water has suggested (see right) that it may be worth exploring a number of standard packages, which could see both totex and ODI incentive packages vary, for instance where maintenance costs were increasing you could see a higher totex sharing rate and more asset health ODI risk within a package, as with higher allowances there is less risk of short term totex underperformance. To balance risk and return for some plans, there could still be a role for local incentives

	Water and wastewater	Water only	Wastewater only
Customers receiving excellent service everyday	C-MeX (residential customer measure of experience) Possibly D-MeX (developer measure of experience) Possibly B-MEX (business customer measure of experience) Possibly R-MEX (retailer measure of experience)	Water supply interruptions Compliance risk index (CRI) measuring drinking water quality compliance Customer contacts about water quality Possibly Event risk index (ERI) measuring impact of drinking water quality problems	Internal sewer flooding External sewer flooding Possibly we could combine the above two as companies are developing a measurement of the impact on customers from flooding
Environmental outcomes	Biodiversity Operational GHG emissions Possibly embedded GHG emissions	Leakage PCC (per capita consumption) Business demand Possibly these could be combined to report distributional input	Pollution incidents Discharge compliance Storm overflows Possibly EPA star rating in addition or instead of above 3 Bathing water quality Possibly river water quality
Asset health and operational resilience	We will decide these following the findings of the UKWIR project 'Future Asset Planning' expect in December 2021.		

Ofwat (November 2021) – PR24 and beyond: Performance Commitments for future price reviews

	Standard	High enhancement spend / environmental obligations	High maintenance spend needed	Long term: Innovative and ambitious
Finance and customer service outperformance	1.00%	1.00%	1.00%	1.00%
Min. Totex outperformance	1.50%	1.50%	3.00%	1.50%
Investment totex outperformance	0.50%	1.50%	1.00%	0.00%
Local ODIs outperformance	0.50%	0.00%	0.00%	1.00%
Resilience ODIs (e.g. supply interruptions) outperformance	0.30%	0.30%	0.00%	0.50%
Environment (e.g. leakage, PCC) outperformance	0.20%	0.20%	0.00%	0.50%
Asset health ODIs outperformance	0.00%	0.00%	0.00%	0.00%
Base return	4.00%	4.00%	4.00%	4.00%
Finance and customer service underperformance	-1.00%	-1.00%	-1.00%	-1.00%
Min. Totex underperformance	-1.50%	-0.50%	0.00%	-2.00%
Investment opex underperformance	-0.50%	-0.50%	0.00%	-0.50%
Local ODI underperformance	-0.50%	0.00%	0.00%	-1.50%
Resilience ODIs (e.g. supply interruptions) underperformance	-0.50%	-1.50%	-1.00%	-1.00%
Environment (e.g. leakage, PCC) underperformance	-1.00%	-2.00%	-2.00%	-1.00%
Asset health ODI underperformance	-1.00%	-1.00%	-3.00%	-1.00%
Upside				
Downside				
	4.00%	4.50%	5.00%	4.50%
	-6.00%	-6.50%	-7.00%	-8.00%

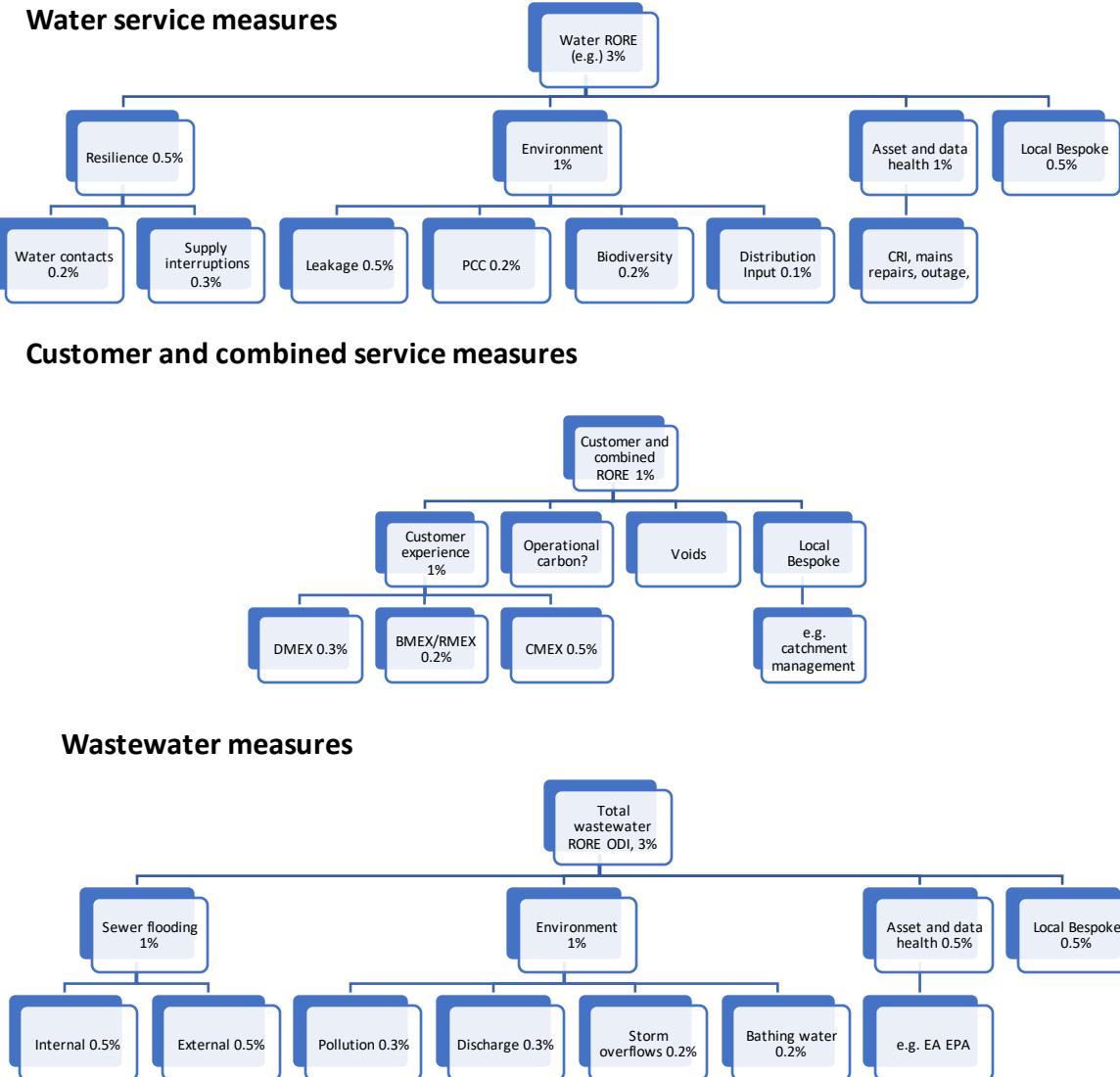
6. If we can base the design of outcome incentives on a “top down” total amount of Return on Regulated Equity (RoRE) that we want to allocate to ODIs, we can test with customers and stakeholders how this balance of risk and return compares to their views. There are two main benefits. First, we have simpler incentives research with customers. Second, we may avoid the risk that in standardising, we merely solve standardisation in one aspect (e.g. service levels and ODI incentive rates), but still end up with risk and return that even on a standard, notional company, basis.

This approach would leave company plans, and trade-offs to justify any bespoke ODIs, informed by company research. It may be relatively easy to standardise or centralise such top down research. By undertaking a top down allocation exercise through customer research (which we explore how this can be done next). We illustrate here what the outcome may look like for the water service, customer experience and combined measures, and separately for wastewater measures. Customer experience measures are included as, from a customer perspective, this is part of the package of incentives that drives companies to deliver for them.

We leave space in this example for asset health measures (including the DWI measures such as CRI, plus mains repairs) as this there is a question as to whether customer research should consider asset health as a single topic, rather than breaking down into component metrics. This rather depends on what the metrics are, and from a risk perspective whether they represent benchmarks based on past performance, or are forward looking predictors based on future factors that affect the performance of those assets.

If ODIs were set top down on a standard % of RoRE, standardised or centralised research could be used to

- Test acceptability of scenarios / plans
- To inform allocation of ODI/RoRE value between measures
- Test design of ODIs to allocate value in support of risk and return balance



7. But how do we make judgements over performance and service levels, particularly P10 and P90 ranges?

One of the key things necessary to reach a robust incentive design is to apply the incentive rates over an appropriate band of performance. One of the key challenges with stated preference surveys is that there may be a budget constraint – a sense that WTP does not vary in a linear manner with different changes in performance. The interaction between budget constraints, the differences between WTA and WTP, and the challenges of scope may in any case be difficult to untangle.

Customer views are unlikely to be linear, and it may vary between metrics. For leakage and asset health metrics, variation in weather may be a factor that means customers do not want incentives applied over a very narrow band of performance. Equally customers may not want exceptionally bad weather to result in penalties so hard that it affects long term investment. This is a challenge that affects both ODI rate research based on “WTP type” approaches and in allocating a top down “RoRE” range to incentives.

The RoRE concept in risk and return requires assumptions on P10 and P90 ranges – the performance expected 80% of the time, and requires a central P50 estimate to be established. It also requires some assumption of a distribution of likely performance (e.g. normal, triangular, even, a long risk tail reflecting low probability/high consequence events) or otherwise within that range.

Incentives also apply outside of the P10 and P90 ranges, but are generally harder to predict – incentive design requires a collar and cap to apply to avoid excessive risk, and also to reflect a level of performance at which the regulatory may need to intervene to protect customers (and in some circumstances for external factors both investors and customers).

In theory for ODI rate research, if we have enough experiment package levels, we can discover customer views on incentives, including where caps and collars are set for P10 and P90. In practice

this seems to be unlikely. The main areas where such issues arise are on asset health, leakage and other aspects where we require incentives, because we are uncertain that the past will necessarily be like the future, with risks such as climate change inevitably uncertain.

The main issues with understanding ranges of performance can be driven by a range of factors, which are all standard regulation and water sector challenges (ie are not really driven by this problem), and particularly relate to the challenges in understanding asset health / resilience and whether past trends are useful in forecasting the future. e.g.

- Uncertainty of variability – e.g. weather, impact of performance improvements
- Uncertainty of external influences i.e. will be the past be like the future?
- Uncertainty of impact i.e. between companies
- Information asymmetry between companies and regulators when presenting risk and performance information
- Attribution of cause and effect (what is inside and outside of management control in the short, medium and long-term is not definitive).

If we can solve the issue of what the service level range is to apply the RoRE range to, then we can simplify the approach and involve some of the risk and return challenges in using bottom up ODI rate research to undertake incentive design. It also may avoid the limitation of the number of performance measures you can include in such research.

We think that outperformance and underperformance incentives are different in terms of the challenges

For outperformance, the design issues appear to be relatively straightforward:

- For ODIs there is less issue with setting P90 ranges, which are inevitably uncertain with forecasting future performance.
- For ODIs where outperformance is appropriate (which excludes most asset health measures where uncertain variation such as weather has the greatest impact), then the performance commitment for one period will normally be in the context of a long-term plan for improving performance. Without this, there is less case for outperformance incentives.
- If this is sector wide performance commitment level, or the five year performance change is standardised, then the ODI change can be standardised (e.g. if the RoRE allowance for leakage is 0.2%, and the increment beyond a performance commitment is for a 10% reduction, then the outperformance incentive 1% is 0.02% RoRE, with a outperformance cap at 10%

For underperformance, setting service level ranges is more complex:

- If service levels and service changes are standardised by Ofwat, then the gap between Performance Commitment and an underperformance collar level can also be standardised.
- It may be that this standardisation should vary by scenario, and also vary with the totex sharing rate.
- Past performance levels also provide an approach, and for asset health measures such

standard interventions based on a mix of company and industry analysis by Ofwat were carried out.

- Experience of PR19 suggests technical approaches to calculated P10 levels are likely to remain subjective, in part because the past only plays a partial role in helping to explain the future – otherwise we would have less need for outcome incentives in the way the water sector finds beneficial to use them.
- One approach is for tiered incentives (which was used for leakage enhancement investment at PR19) – a standard level of P10 range of underperformance for an indicator (as with CRI and supply interruptions), which could vary by scenario or asset management evidence (as with mains repairs).
- The limitations with P10 estimates is they rely on an assumption of normality, which may not apply to individual metrics. There is a risk that systemic risk from weather impacts can apply to a range of metrics, and create a long tail risk.
- Many companies have used Monte Carlo simulation to test the interaction of different ODIs. However, these tools will effectively remain judgements in inputs and how they are used. Ultimately, with no right or wrong answer, we may find that the simplification of defining a set amount of RoRE to put at risk and applying this across a standard range of historic variation for a metric will be just as good as a more complex outcome.
- Below P10 levels of performance, this could either vary by a standard scenario, or be applied to service metrics where the company has a higher historic variation in performance –as an indication of past investment funded by customers.

Asset health ODIs – specific considerations

Marginal benefit-based

- Challenge with using WTP where impacts are indirect or long-term
- Could infer marginal benefit from relationship with customer-facing PCs

Cost-based

- Would fully recover incremental costs of improvements. But what if $MC > MB$ or $MB > MC$?
- Quality of marginal cost data at PR19

Top-down allocations

- How much % RoRE for asset health and over what performance range?
- Could inform balance of overall packages, including where to set caps and collars?



Breakout group questions

What approach should we use? More broadly and for asset health ODIs...

1. If we were to use a **marginal benefits approach** how could we understand the relationships between asset health and service failures over time?
2. If we were to use a **costs approach** to set ODI rates how could we collect consistent information across companies?
3. If we were to use a **top-down approach** how could we decide the performance range in a consistent way between companies?

Look ahead

We suggest the following dates and subjects for **working groups**, although this is subject to change.

13 January	ODIs – Approach for rates that are not based directly on customer evidence
10 February	ODI risk: caps and collars
10 March	Operational resilience and asset health PCs

We intend to publish a **discussion paper on ODI rates** (including enhanced ODIs) in January/February. We will consider wider issues ahead of the draft methodology in June/July 2022.

