Ofwat
Centre City Tower
7 Hill Street
Birmingham
B5 4UA

Strategy & Regulation

Name Nick Fincham
Phone 0203 577 4989
E-Mail nick.fincham@thameswater.co.uk

31 January 2017

Dear Ofwat

<u>Consultation on the outcomes framework for PR19</u>

Thank you for inviting us to respond to your consultation on the outcomes framework for PR19, including the new customer experience measure.[1] We provide detailed responses to the consultation papers and the key consultation questions in Appendix 1. For ease of reference, we set out the most salient points below.

1. Common performance measures

Ofwat needs to exercise its regulatory judgement in comparing companies' performance. We support the use of common performance measures for customer priorities, but the comparisons of company performance must be robust if they are to provide meaningful incentives that stretch all companies. Each company will have its own efficient costs of delivering committed service levels, due to its unique levels of historical investment and outputs, features of the geographic operational area, customer preferences and measurement practices. It follows that each company should target stretching levels of service that take these factors into account. This approach would provide a better result for customers than a one-size-fits-all mechanistic approach that does not take account of these individual factors.

2. Combining incentives, allowed returns and cost assessment

Ofwat should consider performance commitments and incentives in combination with allowed returns and cost assessment, so that the notional efficient company is expected to earn its cost of capital. This could be achieved through appropriate calibration of the combination of the allowed WACC, outcome delivery incentives ("ODIs"), totex allowance and totex incentives. An inappropriate combination, for example where the company is allowed a lower base WACC with ODIs set such that additional returns are earned only for upper quartile

[1] Ofwat, "A consultation on the outcomes framework", 29 November 2016.

or frontier performance on outcomes, would mean that the notional efficient company would not be financeable. This would erode investor confidence and increase the cost of capital, to the detriment of customers. A balanced approach to risk and reward will help promote trust and confidence while also strengthening incentives to deliver the services that are important to customers.

3. **Resilience**

Resilience is a complex issue. The growing challenges facing the sector mean that companies' past experience may not be good predictors of resilience in future. Companies and other stakeholders need to commit to a broad range of activities to understand and improve resilience.

We support the UKWIR and Water UK studies into potential measures of resilience, particularly given there is no current consensus on how to measure resilience, and limited historical information to forecast performance in planning for and responding to events. However, instead of focusing on developing a common measure to compare companies, Ofwat should encourage companies to develop and propose innovative new resilience measures and incentives that demonstrate the benefits to customers of company plans. The outputs from Ofwat's Water and Wastewater Resilience Action Group should help here.

4. **Transparency**

We agree with Ofwat's proposals on transparency. Greater visibility of comparative performance information should mitigate the need for Ofwat to intervene on common outcomes. Performance commitments and incentives should be more transparent, to promote trust and confidence, and minimise the regulatory burden for interim determinations of in-period ODIs. We agree with Ofwat that scheme-specific ODIs will be appropriate in cases where these provide better protection for customers than outcomes-based performance commitments and incentives.

5. **Customer service measure**

We support the review of the service incentive mechanism ("SIM") for PR19. The convergence and subsequent levelling off of SIM scores since PR09 suggests that most of the benefits of catch-up have already been achieved. For AMP6, the convergence of company scores can be recognised within the current methodology by setting the range of rewards and penalties so that material performance differences provide proportionate impacts on prices to customers, for example by retaining the incentive rates per SIM point improvement from AMP5. For the future customer service measure, we consider this would be improved if it:

- is based on the experiences of a wide range of customers, not just a small subset (just 200 customers) that happened to contact companies at specific times;
- provides clear and separable incentives for the wholesale price controls and retail household functions;
- ensures survey scores are comparable between companies, for example controlling for different relative levels of customer affluence in their regions (so that customers in

less affluent regions do not face increased prices because they give materially better scores for any given level of performance than customers in more affluent regions);

- is more closely related to the benefits to customers, within the overall package of outcomes delivery incentives;
- reflects a broader range of channels that customers use to contact companies; and
- focuses on measuring company performance over time, rather than comparing companies, given the need to encourage innovation in customer engagement.

We hope these responses are helpful. If you would like to discuss them further, please do not hesitate to get in touch.

Yours faithfully

Nick Fincham
Director of Strategy & Regulation

Appendix 1 – Responses to the consultation on the outcomes framework for PR19

In this appendix, we provide views on the ten consultation questions and the additional questions raised in the consultation documents.

Making performance commitments more stretching

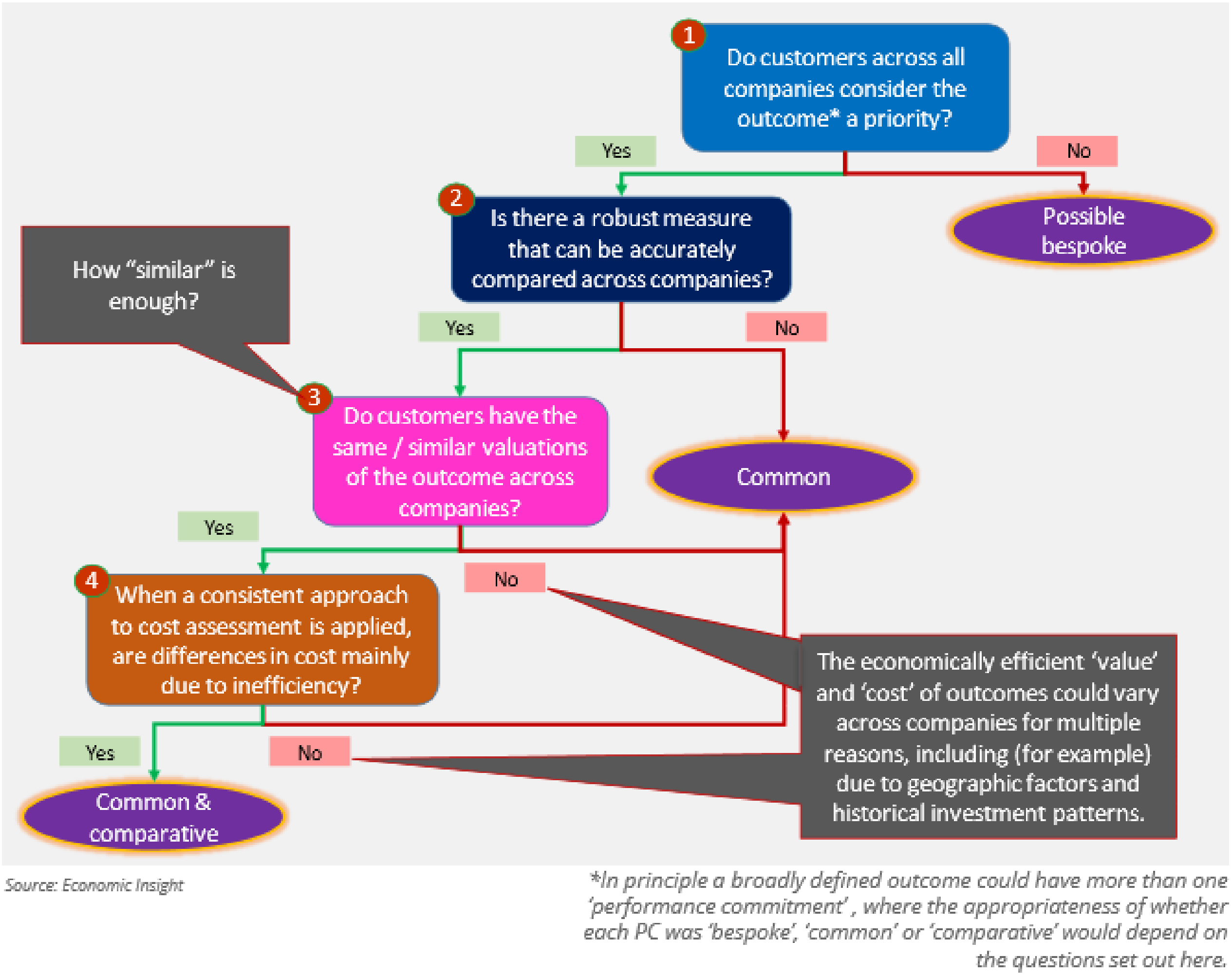
1. What is your view on the use of improved information, including comparative performance information, to make performance commitments more stretching?

We agree that improved comparative performance information will be important for customer engagement and for challenging plans. This information will inform customer preferences to deliver the service levels that customers want based on the efficient costs of achieving those levels of performance, which will improve allocative efficiency. However, there are three reasons why we do not consider that the comparative information can be used directly to set common service level commitments that are stretching for all companies. It will be important that Ofwat applies its regulatory judgement to take account of these factors.

First, to be effective, the comparisons of company performance must be robust and meaningful. The efficient costs of delivering different service levels will vary by company due to varying levels of historical investment and outputs, features of the operational area (e.g. urban or rural density) and differences in customer preferences. This was supported by the findings of a report from Economic Insight,[2] which sets out the necessary conditions for comparative performance commitments, shown in Figure 1.

[2] Economic Insight, "The approach to outcomes at PR19", October 2016, http://bit.ly/2iwL8eG.

Figure 1 – Proposed approach for selecting appropriate types of Performance Commitments



Source: Economic Insight, The approach to outcomes at PR19, October 2016

Second, there will be differences in how companies measure performance, having applied different methodologies and approaches in the past. The initial work through Water UK reviewing company methodologies for a small core set of measures identified a number of material differences. The process for setting comparative targets should seek to adjust for these measurement differences, as otherwise some company targets will not be sufficiently stretching, while other company targets may be unrealistic.

The Water UK review of measures has identified that common definitions for measures is necessary but not sufficient for the performance measures to be truly comparable. The details of how companies measure performance (e.g. interpretation of guidance, direct measurement versus approximation) reflect historical practice and these performance measures have not been subject to horizontal audits previously to check consistency. We support efforts to make the data more comparable, but the effort and time to resolve differences means that the framework should be flexible to recognise that there may be remaining material differences in measurement. This means it would be useful for companies to report their methodology and data confidence levels with any comparative data, to provide context for comparisons.

Third, the measures should reflect better and worse performance by the company, not local circumstances and external factors outside company control. For example, the individual nature of water sources means companies' activities under the Abstraction Incentive Mechanism ("AIM") cannot be compared directly (AIM does not even apply to all companies). Hence there is potential to confuse customers and misrepresent how companies have performed.

The differences in companies' efficient costs and performance due to historical investment and outputs, features of the geographical operational area and customer preferences should lead to different service targets for each company. A mechanistic top-down approach that does not reflect these factors would provide a worse result for customers.

Comparative information can be useful to challenge companies to justify why their performance commitments are stretching. It will be important for Ofwat to apply regulatory judgements about any comparisons so that the performance commitments are truly stretching for all companies, and do not lead to inappropriate performance commitments or incentives that would not be in the interest of customers.

2. What is your view on the common performance commitments we are suggesting for PR19?

We consider that common performance measures will be appropriate where a broad range of customers across all companies consider the outcome a priority. These common performance commitments should have standard definitions, recognising that there may still be differences in how companies measure performance against these definitions.

The proposal for ten common performance commitments ("PCs") is a substantial change from the approach at PR14. We consider that these common measures should be a starting point, from which companies can propose alternative or complementary measures where these are driven by customer research, reflect innovations in how companies measure performance, or are needed to provide better protection for customers. Otherwise, there is a risk that better and more innovative measures are excluded where they would 'double-count' PCs and incentives.

We offer specific comments on each of the proposed common performance commitments below.

New customer service measure

We support the development of a new customer experience measure for PR19, and set out further views in response to question 6 below.

Water quality compliance

We agree that the choice of measure should be informed by the measures that companies will need to report to the DWI in the future. Water quality contacts may be the least suitable measure, as other measures reported to the DWI are likely to better reflect the risk, frequency and severity of water quality failures.

Customer water supply interruptions

We support the use of a common measure, though we will need to test with our customers whether interruptions greater than 3 hours is preferred to alternative definitions.

Our insight to date suggests that customers value planned and unplanned interruptions and events of different durations differently. Moving to a common measure, as defined by the Water UK group for interruptions greater than 3 hours and for pressure below 3 metres would require upgrades to systems, additional resources and re-briefing/training, which would incur additional costs. A move to triggering an interruption at pressures below 3 metres would represent a deterioration in service for our customers, as we currently use 5 metres as the point to trigger an interruption. This means that less customers would be treated as having a problem with an interrupted supply.

Water distribution input (or leakage and per capita consumption)

We do not consider that water distribution input would be a good substitute for leakage as a customer demand measure.

Leakage is a priority to customers and stakeholders, and the importance they attach to it reflects the level of trust they have in water companies. Setting a performance commitment for distribution input would reduce the transparency around companies' performance on leakage, metering and other demand management measures, which could erode that trust. Distribution input can vary significantly over a five-year period due to a number of local and region-specific factors unrelated to efficiency (e.g. population growth and weather). This would make comparisons of improvements more difficult and would reduce transparency around company performance on key measures within company control, such as leakage.

Distribution input will also not be a particularly good measure of resilience, as this requires a view on the balance of supply and demand.

Abstraction Incentive Mechanism ("AIM")

The experience so far from AMP6 is that AIM needs careful design to provide useful incentives. As such we have significant reservations about using it as a comparator of company performance, recognising that further work will need to be undertaken to ensure that such reporting is meaningful and informative, before AIM is included as a common PC.

In our contributions to the AIM Taskforce, we described the challenges in using AIM to give a fair and accurate comparison of water companies' performance. For example, it will be a challenge to represent the many local differences between companies in the number and nature of proposed abstractions, the availability of alternative sources and the associated network, and outage constraints. Historically some companies have already invested in significant environmental improvement programmes where it has been cost beneficial to do so and therefore opportunities to make further improvements will be limited. In these circumstances using AIM as a company comparative measure would not represent a fair and robust comparison between water companies (particularly against those who have not yet made similar environmental improvements) and could misrepresent relative performance. Furthermore, seven water companies did not propose AIM sites so a comparison would only be partial at best.

As a possible alternative, it may be more meaningful for the Environment Agency to consider the publication of a measure reflecting the extent to which company performance on abstraction reduction and mitigation is delivering National Environment Programme ("NEP") requirements contributing to agreed Water Framework Directive ("WFD") objectives in a water body. Objectives based on WFD status or Restoring Sustainable Abstraction ("RSA") could be set relative to what is cost beneficial to achieve in a local catchment and therefore assessment of contributions towards their delivery may surmount some of the reservations about using AIM as a comparator of performance.

Customer property sewer flooding

We agree this is a customer priority. PCs should reflect customer preferences and company history and geographical features. However, we consider that as this measure includes both hydraulic and other causes of incidents, there could be the unintended consequence of incentivising companies to address only other causes of flooding, as this requires less totex than hydraulic solutionsto maintain or improve performance. Companies may need to consider additional flooding measures to complement this common PC to act in customers' best interests.

Wastewater pollution incidents

We agree this is a customer priority. PCs should reflect customer preferences and company history and geographical features.

Asset health water – pipe bursts

We are developing new techniques to help improve the measurement of asset health, although we recognise that the number of pipe bursts is likely to remain a useful measure as part of a basket of asset health measures in the short-term. However, pipe bursts are not directly comparable between companies, as they are a function of age and condition of mains, customer density, materials, diameter, ground conditions and weather, as well as performance. Companies have targeted different service levels in the past to reflect these factors, as demonstrated by serviceability targets (see further details below in response to question 3).

Asset health wastewater – sewer collapses

We are developing new techniques to help improve the measurement of asset health, although we recognise that the number of sewer collapses is likely to remain a useful measure as part of a basket of asset health measures in the short-term. However, sewer collapses are not directly comparable between companies as they are a function of factors both inside and outside company control. Companies have targeted different service levels in the past to reflect these factors, as demonstrated by serviceability targets (see further details below in response to question 3).

An important factor to consider is the impact of the adoption of transferred private sewers in 2011, which means that the historical rate of sewer collapses will not necessarily reflect the future.

New resilience measure(s)

Whether this should form part of the common PCs will depend on the measures developed through the Water UK and UKWIR projects, and any outputs from Ofwat's Water and Wastewater Resilience Action Group. However, our initial view is that resilience is a broad area requiring companies to develop innovative new ways of measuring protection against future risks over the longer-term, in consultation with customers and stakeholders. Therefore, any common measures will need to be carefully evaluated before being proposed as common measures.

One of the challenges will be defining measures in sufficient time for companies to engage with customers and CCGs and to reflect measures in the package of performance commitments, which are part of submitting a high quality business plan. We would therefore welcome early clarity on the resilience measures (and other performance commitments still to be defined) and would suggest that Ofwat shares the outputs from its Water and Wastewater Resilience Action Group as soon as possible.

Additional question: Should the common performance commitments have in-period ODIs attached to them or should this be a decision for the company following engagement with its customers?

We think it should be for companies, in consultation with their customers and CCGs, to propose whether in-period ODIs are appropriate. This may vary by ODI depending on customer preferences.

3. What is your view on how we might apply comparative assessments at PR19?

As set out in response to question 1, we consider that common performance measures will be appropriate where a broad range of customers across all companies consider the outcome a priority, but that common service level commitments will not be in the interest of customers.

We consider that performance commitment levels should be based on customer research and informed by the efficient cost to companies of achieving different levels of service compared with the value to customers of that change in service level. Companies should then be required to demonstrate to Ofwat, CCGs and stakeholders that these performance commitments reflect stretching service levels that provide the best deal for customers when considering cost and performance in the round. This could be achieved in various ways, including benchmarking against historical performance and relevant comparators, using a range of innovative research and engagement methods to understand customer preferences, third party assurance of the planning process, and reviews by the CCG and other stakeholders.

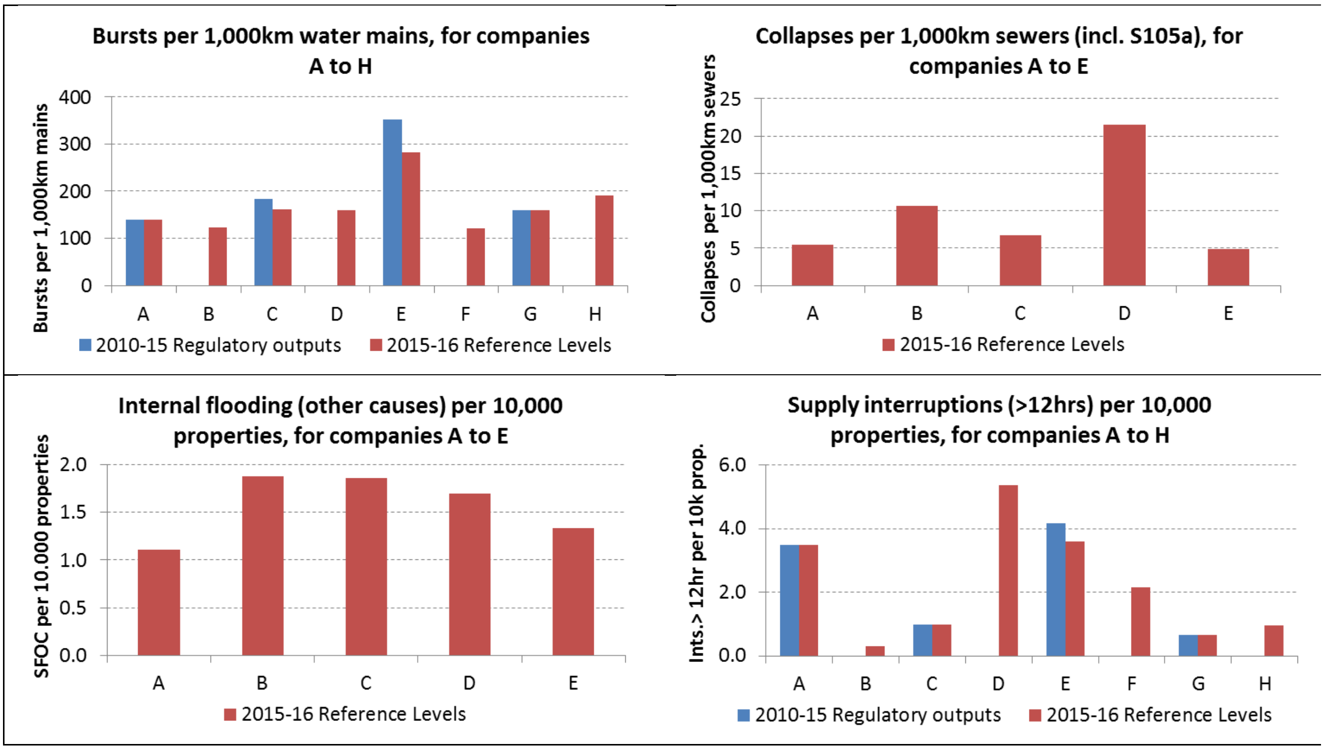
Commitments set at the same levels without reflecting the differences in customer preferences and differences in the efficient cost of achieving particular service levels will produce a worse result for customers, either because the performance targets are not sufficiently stretching or the costs of achieving the performance commitments are disproportionate and more than the amount customers have indicated they are willing to pay.

Companies may be able to demonstrate some of these differences using comparative data available from annual reporting, Discover Water and cost assessment submissions. However, the burden of evidence should be appropriate, recognising that many relevant indicators have not been collected in the past on a comparable basis to demonstrate all of these differences.

We consider therefore that Ofwat will need to apply its regulatory judgement on top of any comparisons of company performance, rather than setting targets in a mechanistic way that does not stretch all companies or leads to inappropriate performance commitment levels. This will help ensure performance targets are sufficiently stretching for the totex allowed in the final determinations.

At previous price reviews at PR09 and PR14, the asset health indices, for example, reflected comparisons of targets with historical trends, rather than trying to compare companies over time. This has resulted in very different targets between companies. For example, some companies have normalised performance targets that is more than double the performance targets for other companies for mains bursts, with even larger differences for collapses and supply interruptions. This is shown in Figure 2.

Figure 2 – Comparison of companies' AMP5 serviceability and AMP6 asset health targets (normalised)



Note: companies anonymised and represented by letters A to H
Source: Thames Water analysis of Ofwat information about FD PC and ODI database, 2016 and Ofwat, cost assessment submissions

The justification for comparing trends over time rather than comparing companies was summarised in Ofwat's serviceability guidance for PR09.[3]

> *"Trends in the metrics are more important than the actual values. Planned step changes in the metrics and related funding are determined through the periodic review process.*
>
> *We recognise that circumstances and data capture practices do vary across companies. We may sometimes find it useful to compare and contrast performance in terms of absolute values across the industry, whilst recognising this has limitations. Such comparisons may inform our judgements about the regulatory action that may flow from a deteriorating serviceability assessment.*
>
> *… We exercise judgement over the value of best historic levels of our serviceability indicators, for example by taking into account the consistency and variability of the data, which might otherwise yield an unrealistic minimum for reference purposes."*

Ofwat states that it may be counter-intuitive if companies receive incentive rewards and penalties from different levels of performance.[4] In practice, however, there would be circumstances where

[3] Ofwat, RD 15/06: Assessing serviceability, annex 1, page 3, October 2006 (available at: http://webarchive.nationalarchives.gov.uk/20150624091829/http://ofwat.gov.uk/regulating/casework/reporting/ltr_rd1506_assessservicbilty).

[4] Ofwat, "A consultation on the outcomes framework for PR19, Appendix 1 – Making performance commitments more stretching", November 2016, page 11.

this would be in the interest of customers. For example, if a particular performance level reflected an improvement on current performance for company A, but a deterioration for company B, then company A should earn a return on the additional cost of achieving the higher performance (assuming this cost did not form part of base allowed totex), and company B should be penalised to compensate customers for the lower performance (assuming totex baselines reflect the efficient costs of these companies' current, different levels of performance).

Applying a restriction that rewards and penalties should be for the same levels of performance does not reflect the different historical and geographical circumstances of each company, so would lead to a very different scope of incentive rewards and penalties for each company, unrelated to efficiency and performance. This would be unfair to customers where there is a high cost for the company to achieve further performance improvements and where customers are unwilling to pay for them. In many instances significant industry innovation is required to make investment cost beneficial for customers, and companies should be encouraged to innovate by earning a return for this improvement in performance.

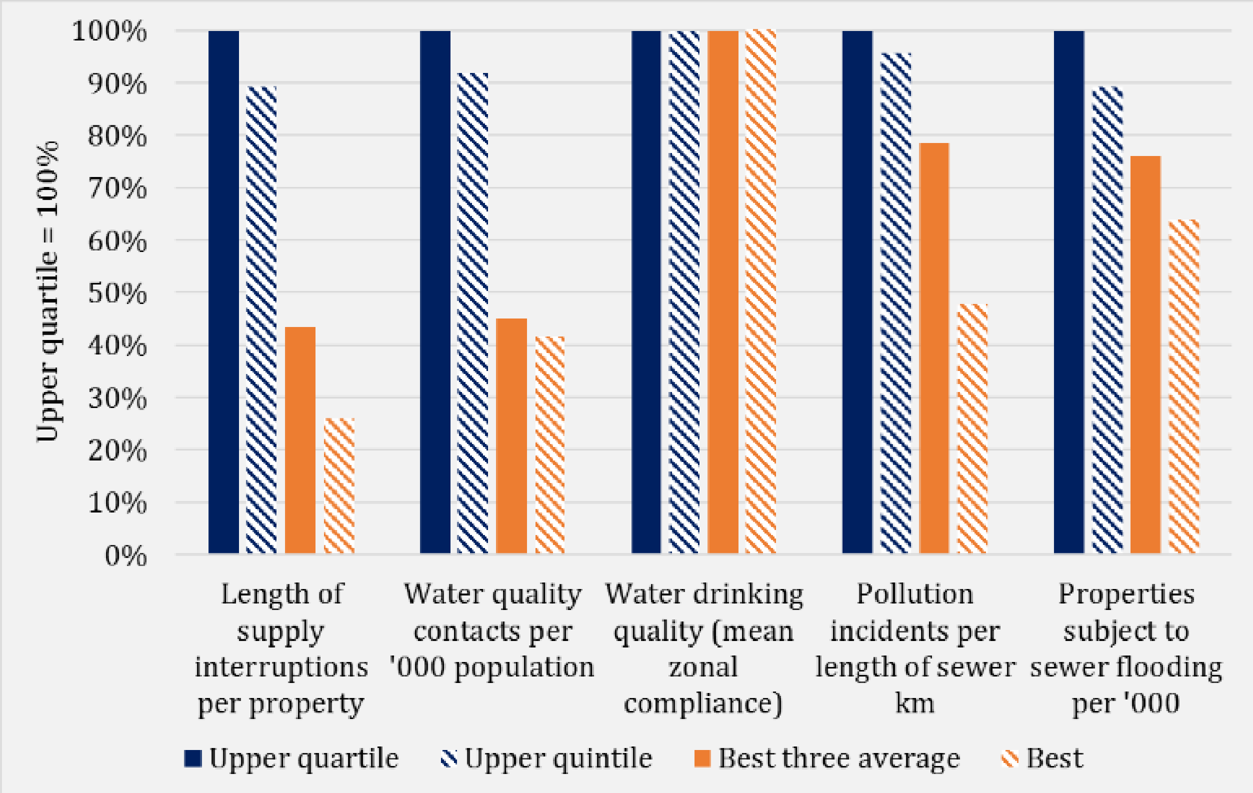
Notwithstanding our views on common performance commitment levels set out above, we provide views below on the specific design questions set out by Ofwat.

What type of assessment?

As stated above, our view is that comparative targets should not be set without first considering customer preferences, historical investment and geographical differences.

In the event that comparative assessments are made, the report from Economic Insight sets out that the upper quartile or upper quintile are less stretching benchmarks than the other alternatives proposed by Ofwat (by definition), but may provide more robust and stable benchmarks than the other choices, while still stretching companies. This is illustrated in Figure 3 below, which compares the performance levels (as index values) for the PR14 common measures at the different options for benchmarks set out by Ofwat.

Figure 3 – Performance levels from Ofwat's benchmark options



Source: Economic Insight, The approach to outcomes at PR19, October 2016

What performance data is used?

An advantage of using historical information is that it reflects actual performance, rather than forecasts which can be over or under estimates.

We consider that static benchmarks will provide stronger and more effective incentives than dynamic benchmarks because:

- Static targets will provide the strongest incentives for technical efficiency improvements within a five-year control period, where changes in customer preferences and technology are slow to happen.[5] This will be more applicable to wholesale activities than to retail activities.
- Dynamic targets are inherently uncertain, which could encourage shorter-term planning rather than longer-term planning where the returns from longer-term solutions are too uncertain. This short-term behaviour could therefore have the unintended consequence of leading to higher bills over the long-term.
- Dynamic performance targets will be difficult to set in practice within a five year period. For example, year-on-year upper quartile performance has been very volatile, making it difficult to identify genuine differences due to performance from factors outside company control (e.g. due to weather), which can be used to adjust performance commitments dynamically.

[5] Economic Insight, "The approach to outcomes at PR19", October 2016.

Annual data or averaging?

We consider that averaged data is likely to be preferable, to avoid any benchmarks being based on a single, potentially unrepresentative, year (e.g. if weather conditions were abnormally favourable or severe). It would still be appropriate to sense-check the benchmarks against the most recent year, to avoid any step-changes.

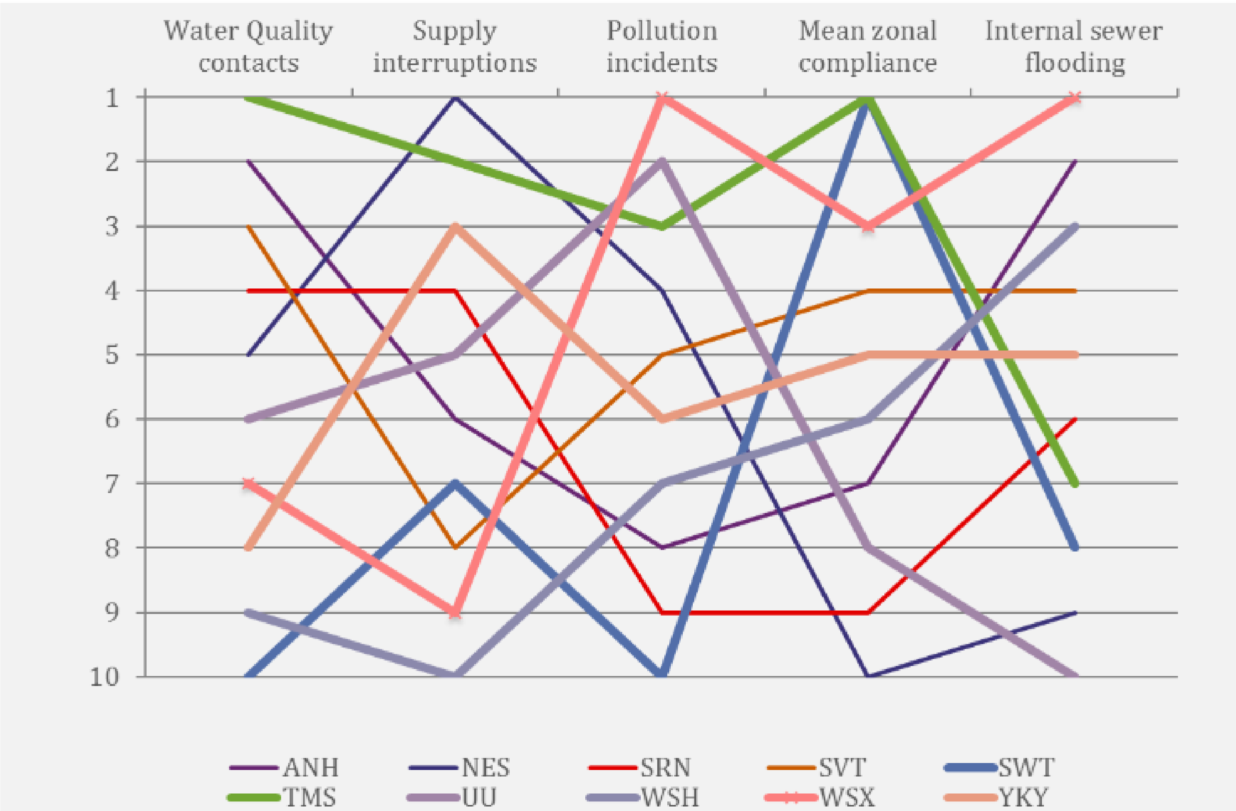
Service measures display a degree of natural volatility, and are often affected by shorter-term influences such as the weather, customer behaviour, company operational practices, as well as defective assets. Performance will therefore not always indicate a change in underlying service. Therefore a review will always be required to assess whether data is representative.

Annual targets are probably more transparent than multi-year averages, though this will depend on whether deadbands are used. For example, either deadbands or averaging of performance can be used to control, at least in part, for the uncertainty from the impact of external events.

What is comparative data applied to?

We agree that setting targets for individual measures will be more transparent and provide clearer incentives. However, to achieve the upper quartile for each individual measure, a company would need to operate far beyond the industry upper quartile, when taking account of trade-offs. This was illustrated by Economic Insight in Figure 4, which shows that no company was consistently an upper quartile performer, even when only considering the five common measures from PR14.

Figure 4 – Relative company rankings by PR14 performance measures



Source: Economic Insight, The approach to outcomes at PR19, October 2016

Economic Insight also analysed quality of service performance in other, competitive, industries. Here it found that, even firms that performed at the 'frontier' overall, had varying performance levels at a less aggregate level, consistent with trade-offs existing between different aspects of performance.

This suggests that the impacts of this portfolio effect need to be considered in setting targets for individual measures, even if a basket approach is not used.

Should comparative data be applied to the PC or ODI?

Whether any comparative data is used to challenge performance commitment levels or other metrics (e.g. penalty deadbands) will depend on the type of measure. For example, a compliance measure may be set at 100% but PC deadbands are challenged based on comparative performance or historical variability in the performance measure.

What glide-path to the target should there be?

Ofwat states that glide-paths will likely not be required at PR19, as all companies will be expected to provide efficient service levels from the start of the next control period.[6]

In our response to question 1 we provide a number of reasons why the efficient levels of costs and performance will vary by company. These factors should be considered in setting both performance levels and glide-paths, so that these are fair across companies and do not require disproportionate costs to achieve.

Where step changes in performance are required, companies will need to invest or change their way of working to achieve these service levels. PCs may require medium or long-term structural changes to achieve, so short-term improvements are not achievable or completely within company control. Where improvements will take a period of time, glide-paths to stretching performance levels are likely to remain appropriate and companies should set out why these are appropriate in their business plans.

Adjustments for company specific factors?

Company specific factors enable targets to be set that reflect external factors outside company control, different efficient costs of achieving service changes and differences in how companies measure performance, which are unrelated to relative (in)efficiency. Company specific factors are important, as without them PCs will not be equally stretching across companies.

We consider there should be a clear process for assessing why there are differences in companies' service levels. However, this will require regulatory judgement, rather than a purely mechanistic approach, where cross-industry evidence is not available to show the full impact of company differences.

Economic Insight suggested that it could be appropriate to set tolerance ranges around performance that allow for uncertainty in determining setting performance levels.[7]

[6] Ofwat, "A consultation on the outcomes framework for PR19, Appendix 1 – Making performance commitments more stretching", November 2016, page 33.
[7] Economic Insight, "The approach to outcomes at PR19", October 2016.

4. To what extent do you agree with our proposed approach to leakage performance commitments for PR19?

We agree that leakage performance will not be directly comparable across companies.

Companies' performance commitments on leakage should be the result of the Water Resource Management Plan ("WRMP") process. The planning guidelines published by the Environment Agency set out clearly defined expectations and challenges for leakage and demand management reduction programmes at PR19.[8] These ensure that the sustainable economic level of leakage is fully integrated with the WRMP and is challenged to make sure it reflects longer-term benefits and innovations.

The guidelines show the Government has a strong focus on taking a long-term, strategic approach to water resource planning for leakage reduction and demand reduction that represents best value to customers in the long-term. We therefore consider that any Ofwat challenge to leakage targets should form part of the WRMP process. This will help provide assurance that leakage targets are stretching when taking into account the balance of different supply and demand solutions available. It would also avoid the need to make any late changes which could undermine the WRMP process and duplicate effort at the price review. Ofwat should also ensure that its approach to setting access prices for the bilateral market, particularly whether incumbents can defer leakage reduction schemes due to entry, is consistent with its approach to setting leakage targets.

Adjustments to the totex modelling may be required so that allowed costs reflect the improvements required by each company. This could form part of the special cost factor process.

5. What factors should we take into account in our guidance on setting performance levels for bespoke performance commitments at PR19?

We consider that bespoke PCs should be driven by companies' customers and stakeholders, as well as regulatory and legal requirements. In general, this should mean that Ofwat does not need to mandate PCs in specific areas and should require minimal guidance over and above the proposed guidance around transparency (see question 10 below).

It would be useful for Ofwat to provide expectations on the financial ODIs associated with bespoke commitments and how these fit within the overall risk-return package. For example, we would welcome guidance on how companies might set ODI rates, deadbands and caps/collars. This would be necessary to achieve consistent links between returns, ODIs and cost assessment.

Additional question: Are there particular areas that all companies should cover with their bespoke performance commitments?

The common PCs proposed by Ofwat cover a number of key service areas. Ofwat has also set out expectations in other areas, such as leakage and vulnerability. Companies also need to comply with legislation and other regulations. In general, we consider companies should be given ownership to design bespoke PCs with their customers and CCGs and it is not necessary for Ofwat to give further guidance on the areas to be covered by bespoke performance commitments.

One area where Ofwat could consider providing guidance is PCs for developer services. We consider that revenue from developer services should be excluded from the price controls at PR19.

[8] Environment Agency, "Leakage in WRMP19", November 2016.

If this is not the case, then it may be appropriate to provide separate incentives on companies around meeting actual growth, to support housing development.

Additional question: Should Ofwat require companies to report against common metrics on vulnerability (e.g. proportion of customers falling into arrears with their bills, or percentage of eligible households receiving a social tariff)?

We agree that companies should consider PCs that demonstrate what they are doing to protect customers in vulnerable circumstances. Definitions of these vulnerable circumstances should be flexible to allow companies to address different situations that may arise and common measures may not be appropriate where companies are taking different approaches.

Additional question: Should companies propose performance commitments and ODIs that relate to specific price controls?

We agree that PCs and ODIs should, where possible, be specific to particular price controls. This may not be straightforward where services are delivered by balancing activities that span more than one price control (e.g. the water supply-demand balance). This includes some of the common measures proposed by Ofwat, where an industry wide approach to allocation to a specific price control would be appropriate.

This may mean that some price controls are more likely to earn rewards, which may need to be reflected in the allowed cost of capital. The framework needs to ensure that changes to expand markets and competition do not blunt the power of incentives, restrict innovation or distort competition. One possible solution would be to assign the ODI rewards and penalties associated with bioresources and water resources to the network-plus controls.

In any cases where it is not practical or in customers' interests to assign incentives to a single price control, the approach to sharing incentives between price controls should be clearly defined up front and be transparent to customers and stakeholders.

6. What is your view on our development of a new customer experience measure for PR19?

We support the review of the service incentive mechanism ("SIM") for PR19. The convergence and subsequent levelling off of SIM scores since PR09 suggests that most of the benefits of catch-up have already been achieved and it is an appropriate time to review the measure.

We consider there are ways to recognise the convergence of company scores when setting the range of rewards and penalties for AMP6 SIM. For example, retaining the incentive rates per SIM point improvement from AMP5 would provide proportionate impacts on prices to customers for material differences in scores, which should reflect tangible performance differences and additional benefits to customers.

The risk of an approach where relatively small differences in scores (unrelated to performance) have disproportionately significant impacts on prices to customers, which do not reflect benefits to customers, is exemplified by Economic Insight's findings that:[9]

> *"none of the water [performance] commitments being statistically significant in explaining the SIM as a whole and its qualitative and quantitative constituents".*

---

[9] Economic Insight, "Future customer service incentive for water", November 2016, page 56.

The prospect and timing of residential retail competition is uncertain and companies' wholesale functions will continue to contribute to customers' perceptions of the quality of service they receive. Therefore, we consider there is a role for future customer service incentives that encourage:

- efficient behaviours, by aligning incentives with customer benefits;
- improvements by all companies, whether good or bad performers;
- consistency in how companies are measured and fit within the outcomes framework;
- transparency for customers and stakeholders; and
- innovation in how customer service improvements are delivered.

As shown by the Economic Insight report commissioned by Ofwat, there are a number of possibilities and trade-offs in designing the future customer service measure. We are doing further analysis to help identify the best options for the specific customer service measures that would best answer the criteria above, which we will share with Ofwat ahead of the PR19 methodology.

At this stage, we provide views below on the design of the future customer service measure and the key questions raised by Ofwat.

The ultimate outcome?

The new customer service measure should complement companies' other performance commitments, including reflecting the results of companies' engagement with customers and the benefits to customers from changing service levels.

This could be achieved by designing a measure and incentive that trades off the customer benefits against the cost of delivering those benefits. Such a company-led approach based on the findings from customer research, would encourage companies to push the frontier forward to achieve dynamic efficiencies to deliver what their customers want.

There is an opportunity to design a more rounded measure of customer service that includes views from a wider range of customers including those who do not contact the company in a specific time period. The surveys of customer satisfaction and trust should enable views from a wider range of customers to be included, though value for money may also be a key component of customer experience. This approach would require multiple questions in any survey, so a balance is needed between the data requirement and the ease of responding to the survey. Moving away from being solely a telephone based survey to include other methods such as SMS or online might improve this.

For any comparative assessment, survey scores must be comparable between companies. In particular, recent work by Deloitte[10] showed that the propensity of customers to award good or bad customer service scores for any given level of service is correlated strongly with affluence. Customers in less affluent regions are shown to give materially better scores for any given level of performance than customers in more affluent regions. This means that customers in less affluent regions will face increased water prices merely because they give higher satisfaction scores for the same level of service as more affluent customers. We do not believe that this is in the interests of customers.

[10] Deloitte, "Customer service incentives: statistical exploration of the water industry SIM", May 2016.

Any survey approach should also provide a robust reflection of company performance by using statistically robust samples on an ongoing basis during AMP7 (e.g. for each quarterly survey). For example, samples could be based on a proportion of total contacts or customers, rather than the same absolute sample size for each company.

Wholesale to retail?

The measure should cover wholesale and retail to reflect full customer service, and should provide clear and separable incentives for water network-plus, wastewater network-plus and retail household functions. More direct incentives improve efficiency and effectiveness, and are flexible to market development.

Separately defining wholesale and retail performance will be challenging, as companies have different operational models and retail has responsibility for operational calls – the current allocation doesn't allow customer service satisfaction to be split. Operational contacts should form part of wholesale if the responsibility and satisfaction with operational contacts are to sit in the same place. Incentives should be aligned with the business that has responsibility for performance.

The survey piloting will need to check that the survey design provides an accurate reflection of company performance for each business. If not, then an appropriate allocation approach will need to be developed.

Beyond water?

The measurement of customer service could consider comparisons beyond the water sector. This could help control for geographical differences, though industry-wide differences would need to be accounted for, so direct comparisons to set benchmarks and incentives are unlikely to be appropriate.

Beyond contacts and complaints?

As set out above, we support the design of a more rounded measure of customer service that includes views from a wider range of customers including those who do not contact the company in a specific time period.

A multi-channel approach?

Measurement of customer service should reflect a broader range of contact channels, to reflect innovations and developments in customer service. For example, this could include self-service (through IVR or online), social media (received as being any contact via message inbox or where the company chooses to respond to the customer, webchat (offered or advertised) and SMS text.

The inclusion of different channels would support the measurement of individual company performance over time. However, comparisons between companies would need to control for the different channels employed and the large number of definition problems that would arise. Otherwise, measures would reflect the channels that companies use rather than customer service performance and might provide skewed incentives on choices of contact channel. This issue is highlighted in Economic Insight's report which shows the wide range of channels that customers use and that work to understand how to summarise customer sentiment from social media is an ongoing issue.[11]

[11] Economic Insight, "Future customer service incentive for water", November 2016.

The role of complaints?

We agree that measurement of complaints will continue to have an important role and so it seems appropriate to retain an appropriate measure.

There may not be a strong case for an additional financial incentive on complaints. Companies already face strong financial incentives to reduce complaints through the impact on satisfaction scores, measurement by stakeholders such as CCWater, the cost savings associated with reducing complaints (companies currently benefit from 100% of cost savings from reduced complaint volumes until the end of AMP6), and the prospect of residential retail competition (as high complaint levels would indicate a higher propensity for losing customers to competitors).

The measurement of complaints would be significantly improved through more rigid and consistent definitions. If a measurement of complaints is retained, the current guidance should be reviewed to ensure greater consistency between companies' reporting of performance.

Incorporate vulnerability?

As set out under question 5, we agree that companies should consider PCs that demonstrate what they are doing to protect customers in vulnerable circumstances. Definitions of these vulnerable circumstances should be flexible to allow companies to address different situations that may arise and common measures may not be appropriate where companies are taking different approaches.

This may mean that a separate element on customer service for customers in vulnerable circumstances is not required.

Beyond end-users?

As set out above, we support the design of a more rounded measure of customer service that includes views from a wider range of customers including those who do not contact the company in a specific time period. This could include customers who come in contact with companies who may not necessarily use water or wastewater services and Developer Services contacts from residential customers.

We note that Water UK's Infrastructure Policy Group is planning to explore a new system for measuring developer customers' views.

Summary

In summary, we consider the future customer service measure would be improved if the measure:

- is based on the experiences of a wide range of customers, not just a small subset (just 200 customers) that happened to contact companies at specific times;
- provides clear and separable incentives for the wholesale price controls and retail household functions;
- ensures survey scores are comparable between companies, in particular so that customers in less affluent regions do not face increased prices because they give materially better scores for any given level of performance than customers in more affluent regions;
- is more closely related to the benefits to customers, within the overall package of outcomes delivery incentives;
- reflects a broader range of channels that customers use to contact companies; and
- focuses on measuring company performance over time, rather than comparing companies, given the need to encourage innovation in customer engagement.

More powerful outcome delivery incentives

7. What is your view on the options for increasing the power of reputational and financial ODIs at PR19?

We welcomed the approach at PR14 of companies proposing outcomes, performance commitments and incentives in line with customer research. This supported the objectives of greater customer engagement and company ownership in business planning.

For PR19, Ofwat is considering strengthening the link between rewards from outcome delivery incentives and the levels of the base return. We support, in principle, the strengthening of incentives for operational performance, though it is essential that the design of the outcomes and incentives framework is fully underpinned and supported by customer and stakeholder engagement.

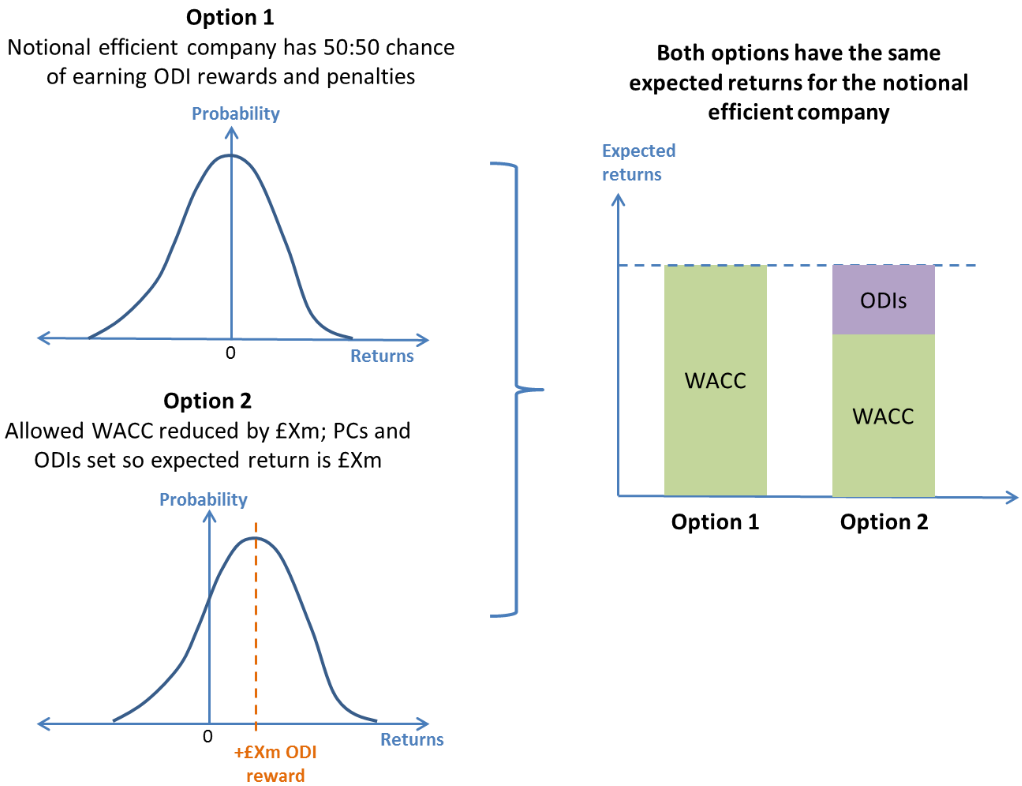
In our response below, we comment separately on: (i) the principle of strengthening ODIs in the context of the overall risk and reward package; and (ii) the specific proposals by Ofwat for strengthening ODIs.

(i) Principle of strengthening ODIs in the context of the overall risk and reward package

The risk and reward package will require careful implementation, so that overall the notional efficient company can earn its cost of capital, either through the allowed WACC in line with historical approaches or through a mix of WACC allowance, ODIs and totex outperformance. This would retain investor confidence in the sector, keeping bills lower in the longer-term.

Figure 3 in Ofwat's outcomes consultation, on how a penalty-only approach can lead to higher bills, is confusing. The diagram would only reflect the principle above where the reduction in the base bill (from £400 to £390) from the reduction in allowed WACC is offset by an ODI reward at the expected level of performance for the notional efficient company. We illustrate this principle in Figure 5.

Figure 5 – Illustration of link between ODI probability function and base returns



Source: Illustration from Thames Water

This means that the package of allowed industry WACC, ODIs and totex incentives should be considered in combination, with ODIs appropriately calibrated to ensure that the incentive properties for operational performance are retained or enhanced, whilst still ensuring that a notional efficient company can earn the overall required notional industry cost of capital. For example, this could be achieved where the base WACC plus expected returns from ODI outperformance and totex outperformance are equal to the overall required notional cost of capital for the notional efficient company.

If inappropriately applied, so that the base allowed WACC and expected outperformance of ODIs and totex is less than the overall required notional cost of capital, a notional efficient company would not be financeable. This would erode investor confidence and increase the cost of capital, to the detriment of customers. This would occur, for example, where the company is allowed a lower base WACC with ODIs set such that additional returns are earned only for stretching outperformance of outcomes that a notionally efficient company only has a 50% change of achieving. In this situation an efficient company, meeting its performance obligations, would be unable to earn the overall required notional industry cost of capital.

In addition, levels of service must be sustainable over the long term. An incentive scheme should not be put in place that incentivises gains in service over the short term at the expense of maintenance or long term asset replacement.

To illustrate the difficulty of an approach where the reduction in base returns is offset by additional ODI rewards at the industry upper quartile level of performance, we calculate that all water and wastewater companies would need to operate at least at the P90 levels of performance for the performance commitments set by Ofwat’s horizontal checks to offset approximately a 0.2%

reduction in WACC.[12] This means having a distribution of ODIs that is much more heavily weighted to rewards than penalties, as shown in Figure 5.

Another important consideration when thinking about the allowed WACC and ODIs together is that the design of the ODI mechanism may itself have an impact on the cost of equity. If ODIs are made more powerful this would increase the volatility of potential revenues earned by companies. Investors would require this additional volatility to be compensated for by a higher cost of equity component of the WACC. Similarly, if a company committed to more stretching ODIs, thus increasing the risk of achieving them, the increase in asset beta means the company would require additional returns. In addition, if investors perceive returns from ODIs as riskier than allowed WACC, they could require a premium on expected returns if stronger incentives are used in place of allowed WACC.

Annual volatility in performance (e.g. due to weather or other factors outside company control) will be managed to ensure companies are financeable on an annual basis. This is especially pertinent if ODI rewards and penalties are incurred within period, as trends different to forecasts could have significant impacts on in-period financeability if not managed carefully.

It is also worth noting that reference data points typically used to inform the cost of capital (such as equity betas) will not reflect any move to more powerful ODIs and the associated increase in volatility of returns. Further work will be required to better understand and quantify the relationship between these variables - it will be critical to properly assess the impact on the cost of equity through adoption of more powerful incentives, any increase to systematic risk arising should be carefully evaluated and added to the cost of equity component of the WACC.

We note that the consultation refers to the potential for a menu-based approach to the cost of equity. Ofwat sought initial views from companies within the October 2016 consultation on the cost of debt. Whilst reference should be made to our detailed response to that consultation,[13] our view remains that we are not convinced that a menu based approach to the cost of equity (as considered by the Essential Services Commission in Victoria) is as good as the risk based approach used by Ofwat in PR14. The menu approach appears to include a number of issues that would need to be dealt with in advance of being introduced and, at this time, it is difficult to see how such an approach could be implemented in a way that is demonstrably fair to investors and customers.

We have commissioned EY to examine different ways in which the allowed WACC, totex incentives and ODIs could be combined together into a coherent, reasonable overall risk and reward package for PR19. We expect this paper to be available on the Water 2020 marketplace for ideas around mid-February 2017 and will highlight this report to Ofwat when it is available.

(ii) Proposals by Ofwat for strengthening ODIs

Ofwat is considering whether a more symmetrical approach to ODI rewards and penalties would help drive better outcomes for customers. There will be a number of important considerations for effective implementation:

[12] This is based on a simple calculation from data available in the published Final Determinations. An aggregation of the P90 rewards published by Ofwat for the horizontal check PCs is around £700m (2012-13 prices). This is roughly equivalent to approximately 0.2% on the forecast closing 2019-20 RCV of £65bn over five years.

[13] Thames Water, "Water 2020: consultation on the approach to the cost of debt for PR19", October 2016.

- Customers and stakeholders will need to support more symmetrical rewards and penalties. For example, ODI rewards need to be shown to be reasonable, rewarding companies for the effort and risk-taking to encourage innovation and frontier-shift, which will benefit customers.
- Increasing the absolute size of rewards and penalties (e.g. by removing caps and collars) is not necessarily effective. Instead, the relevant focus is how performance commitments are set and the probability of earning different rewards and penalties. This will determine the expected return and realistic range of potential returns, which should be the focus for Ofwat, companies and stakeholders in assessing the overall risk and reward package, as companies should be able to earn the returns through innovation that improves services for customers. The P10/P90 approach from PR14 was a move towards this, though more sophisticated analysis should be used at PR19.
- Removal of deadbands could strengthen incentives, but also have the potential to make rewards and penalties more asymmetric where performance commitments are made more stretching.
- The bottom-up formulae to set incentive rates at PR14 were asymmetrical by nature. ODIs could be made more symmetrical with a small tweak to the ODI penalty formula while still encouraging the right behaviours by companies. This is shown in Figure 6.

Figure 6 – Proposed approach to setting more symmetrical incentive rates

As set out in Ofwat's outcomes consultation, Appendix 2 (page 18), ODI rates were calculated as follows at PR14:

$$ODI_{penalty} = incremental\ WTP - p * incremental\ MC$$
$$ODI_{reward} = (1 - p) *\ incremental\ WTP$$

where *WTP* = willingness to pay, *p* = totex sharing rate for customer, *MC* = marginal cost

Where performance commitments have been set so that customers are not willing to pay for further outperformance at the expected cost (i.e. incremental WTP < incremental MC), the ODI penalty rate will be larger than the ODI reward rate by definition.

This is because ODI rates are calibrated so that, where companies overspend to earn rewards, they share the gains with customers. But where companies underspend and earn penalties, customers are held neutral and companies bear the full downside.

The ODI penalty rate can be changed to align with the ODI reward rate, so that companies are still incentivised to avoid underperformance, but the downside is shared in the same way as the upside, i.e.:

$$ODI_{penalty} = (1 - p) *\ incremental\ WTP$$

Source: Thames Water

There will be other practical reasons why ODIs may be focused on penalties rather than rewards. For example, at PR14 some stakeholders expressed strong views that companies should not earn ODI rewards for achieving legal obligations (e.g. around pollution incidents and discharge consents) and the use of penalties (but not rewards) around delay or cancellation of major schemes. These areas will need to be considered by Ofwat and companies in designing more balanced risk and reward packages.

We agree that in-period ODIs and revenue adjustments may provide stronger incentives, though the form of incentive should be underpinned by customer research and likely bill impacts. Any use of in-period ODIs must be balanced with long-term asset stewardship, to ensure they do not incentivise short-term behaviour.

We agree that greater transparency in reporting good and poor performance will strengthen the power of reputational incentives. This could include updates to Discover Water to show company performance against performance commitment targets using a single, simple to understand format.

We agree with Ofwat that gated ODIs should be avoided where possible as these reduce the incidence and therefore strength of incentives.

Better reflecting resilience in outcomes

8. What is your view on our proposals for better reflecting resilience within the outcomes framework?

In general, we support the draft resilience planning principles proposed by Ofwat, as these should reflect how companies are approaching this part of business planning. Enhancing the resilience of services is distinct from maintaining asset health.

Resilience planning is a complex area and involves understanding the resilience of services to customers, not just assets. Resilience includes company actions to avoid events, reduce the impact of events and the ability to recover after an event has happened. Company approaches to planning should be forward-looking and flexible, to consider the additional risks that companies need to plan for, including the underlying quality of assets, which continue to deteriorate with age. By their nature, low frequency events can be difficult to predict from historical failure rates. Therefore as companies' data and understanding of system and asset resilience are improving over time, their ability to predict failure with cost effective intervention will improve.

Companies and their customers will have different levels of resilience and risk appetite. As companies cannot be fully resistant to all events, companies should engage with customers in determining the appropriate levels of resilience and how to most effectively and efficiently achieve this. For example, a sustainable and resilient system could require the rate of asset replacement across the industry to increase as well as alternative methods of delivering resilience. Companies, in engagement with customers, will need to consider whether the short-term bill impacts are acceptable when compared with declining resilience and higher whole-life costs and bills for future generations of customers. There is a balance to be struck between the affordability of increasing investment in resilience, compared with the affordability of measures needed to recover from significant events should they occur, in the absence of such investment.

We agree with the need for measures of resilience. However, the measures should consider companies' understanding of the risk to services and the range of avoidance, response and recovery actions that companies can take to events. The work by UKWIR and Water UK should identify options for useful measures to monitor. However, these should not be used to set common targets with financial incentives that cut across the need for flexible planning by companies to consider the different facets of resilience and how it evolves over time. Actions to improve resilience are likely to be company-specific so would require bespoke approaches to measure their effects. We consider that companies should propose measures that align with their own business plans to drive the right behaviours.

Customer engagement, even where it includes wider economic and environmental benefits, should be balanced with a robust bottom-up process, based on best available information and risk assessment, as these may lead towards different outcomes. The aim should be to balance these approaches appropriately to identify solutions that minimise whole-life costs and provide the best value for money.

We disagree with Ofwat that it will be sufficient for it to impose larger penalties on companies in order to incentivise them to invest to plan for and respond to high-impact, low-frequency events.[14] Companies should be incentivised to put in place the lowest whole-life cost solutions to build resilience into their systems such that service deterioration is avoided where possible. The solutions should be affordable and supported by customers, and complemented by appropriate rewards and penalties, to provide returns on efficient investment and protect customers. It will not be possible for a company to avoid or mitigate all high-impact, low-frequency events, since such events may not be predictable or the investment required may be disproportionate and not supported by customers.

9. What is your view on the options and our preferred approach to asset health outcomes?

We consider that the outcomes framework should encourage companies to continue to propose new and innovative ways to measure asset health, and avoid adopting prescriptive, industry-wide approaches that might discourage this.

We agree with providing additional transparency on asset health reporting, including sub-measures, which we consider will improve trust and confidence in the measurement of asset health.

We agree it would be appropriate for companies, in consultation with customers and stakeholders, to consider rewards for asset health where this provides returns for additional efficient investment that benefits future customers.

Making performance commitments more transparent

10. To what extent do you agree with our proposals for making performance commitments more transparent for customers?

We agree with Ofwat that greater visibility should be provided over comparative performance information, to mitigate the need for Ofwat to intervene on common outcomes. This includes asset health and sub-measures, as set out in response to question 9.

We agree with the principles proposed by Ofwat, though concision should be balanced with completeness and clarity.

[14] Ofwat, “A consultation on the outcomes framework for PR19, Appendix 3 – Better reflecting resilience in outcomes”, November 2016, page 15.

We agree with Ofwat that scheme specific ODIs will be appropriate when this provides better protection for customers and stronger incentives than outcomes based performance commitments and incentives.