United Utilities response to the Ofwat's consultation:

Outcomes framework for PR19

Key issues

United Utilities welcomes the opportunity to contribute to Ofwat's further developments of the outcomes framework for PR19.

In general, we support Ofwat's overall approach of evolution of the Outcomes framework established at the previous price control. We broadly support Ofwat's proposals for common performance measures, and the use of company comparisons, but we note two significant issues that require more attention:

- inconsistent measurement and calculation of measures between companies. The work being undertaken on convergent measures is helping, but full consistency remains challenging in some areas; and
- non-comparable performance levels due to regional factors – it is reasonable to assume that all companies can achieve comparable economic levels of performance (as occurs with leakage), but such approaches are not undertaken for other outcome measures (such as sewer flooding), where a simplistic common target does not reflect the significant difference between companies' ability (and cost) to meet certain performance levels.

We fully support transparency and simplification where possible, as it is in customers interests. However, simplicity should not result in inappropriate or poorly targeted incentives. For example, outcomes should avoid risk of companies focussing on a single narrow target at the expense of performance across a service in total.

Regarding Ofwat's desire to increase the power of incentives, we suggest further consideration be given to the risk of companies facing inappropriate incentives. For example, if incentive values exceed customer support for a particular change in service level, then this could result in customers paying more for service levels that they do not desire. We also note that Ofwat should still have a role to intervene in the event that service levels deviate significantly from expectations, resulting in substantial financial impacts on customers.

We support Ofwat in seeking to develop two new outcome measures, for resilience, and a new measure for customer experience. Our response aims to make a substantial contribution to Ofwat's development in these two areas.

For resilience, we propose that a commonly defined risk assessment methodology (set out in Annex 1) could be used by all companies to derive a suitably consistent measure of resilience risk, based on the "number of customers with high resilience risk". This measure would be assessed net of all company resilience mitigating actions, including response and recovery plans. We have reviewed the measures proposed by other parties, and do not consider that any of these appropriately measures resilience risk. We would be happy to work with Ofwat and other companies to develop this approach further.

For the new measure for customer experience, we support Ofwat's objectives, particularly in terms of; the need to drive further improvements to service levels; adapting to changing customer expectations and behaviours; supporting service innovation; and learning from the design of SIM. We have a number of proposals as to how these objectives can be achieved in practise.

Copyright © United Utilities Water Limited 2016

United Utilities response to consultation on outcomes framework for PR19

Making performance commitments more stretching

Q1: What is your view on the use of improved information, including comparative performance information, to make performance commitments more stretching?

Comparative information

We support the use of comparative information and have fully supported the development of the Discover Water dashboard and intend to use this as part of our engagement with customers, and in interactions with YourVoice, our CCG.

Comparative information is a useful source of data for CCGs and customers during the engagement process to help them understand the relative position of a company, and to support and challenge the company in driving improved levels of performance. However, for comparative data to be useful, there are some further considerations:

- **Consistency in data collection and reporting methodology** – we believe that there should be greater levels of assurance that all companies are collecting the data in a consistent manner. Discussions as part of the convergence and Discover Water work have demonstrated that companies do not all collect their information in the same way and therefore it is likely to be the case that at least some data comparisons reflect differences in reporting methodology (including data capture, as well as definition and application of methodology) rather than variations in performance.
- **Regional factors** – it is to be expected that there will be regional variations in some performance metrics, driven by weather patterns, demographics, geography, differences in inherited assets and other external factors. Any comparison of company performance should be normalised to recognise such factors. We recognise that could be a complex process. As a minimum companies should have the opportunity to explain such differences. If comparable performance is only achievable at varying levels of cost between different regions then customers and CCGs should be aware of this when discussing comparative information.

Whilst comparative information should be used in discussions with customers and stakeholders about performance levels and future targets, it should not be used simplistically to determine performance commitments with no further consideration of whether that would be appropriate for a specific company. Company specific committed performance levels should be driven by the requirements of the customers in an area. The value that customers place on services could vary across the country and in some cases even within regions.

Making targets more stretching

We support further developments in outcomes to explore options that could make performance commitments more stretching at PR19. There are many features of ODI design that contribute to how stretching that outcome is for a company.

- **Use of exclusions** – we recognise that Ofwat considers that allowing for fewer (or no) exceptions in their performance commitment definitions can make performance commitments more stretching. Whilst we agree that companies can, in many cases, reduce the potential impact on their services should such events occur, it is not always possible to mitigate against all external events. On the other hand, appropriately targeted exclusions, such as third party incidents or severe weather (that impact on service beyond that which it is reasonable for companies to be resilient against) can actually allow companies to set more challenging underlying targets. We note that other regulators are content to assign failures to third parties and exclude them from performance data (for example, the EA have historically granted exclusions for issues caused by third parties, and issues caused by extreme weather events such as storm Desmond). Such exclusions should also be commonly defined, e.g. exclusions for extreme weather can be by reference to storm return frequency and/or by published EA river level return data.

Copyright © United Utilities Water Limited 2016

- **Using a baskets of measures** – we agree that using performance against a basket of measures can allow for more stretching performance to be targeted on individual measures, as well as more comprehensively representing performance within a service area (moreso than any single measure). For example, it allows companies to balance performance levels across measures, which reduces the disproportionate risk of penalties on any individual measure. Whilst this approach can be beneficial, we also note that this also adds a certain amount of complexity to the process, and may not be easy for customers to engage with, as is the case for index measures. It should be noted that many of our index measures are also effectively baskets of measures, with the index used simply to provide a normalised basis of reporting. This is a complex issue – for some service areas, too much focus on a single measure may incentivise companies to focus too much on the measure, rather than ensuring that customer expectations are met for all facets of a service.
- **Use of average vs. "in year" performance** – if performance is subject to some natural variation year on year, then some companies may find a tougher target to be more acceptable if performance is measured on a trailing average over a number of years, rather than on a single year's performance. Conversely, using single year performance can be more stretching as it can encourage companies to focus on reducing the impact on their services of single "in year" events. Use of averages also prolongs the impact of single events over multiple years, which can be unduly punitive in the event that the company's underlying performance is on target.
- **Use of deadbands** – deadbands are helpful, and they should be used to allow for natural variation in performance before the company incurs rewards or penalties. The use of deadbands – not just whether or not they are applied, but also both the absolute size, and how symmetrically they are applied – contribute to making underlying targets more stretching.
- **Use of gated ODIs** – gating mechanisms can be used to make rewards (or penalties) contingent on performance on another measure, where there may be relationship between the two measures, for example gating a leakage ODI reward contingent on performance on security of supply performance commitment. This can have the effect of making rewards more difficult to achieve if the company has only performed in one area (although conversely companies could also avoid penalties if performance elsewhere is good). Gated ODIs would increase the level of complexity within calculations, and may be confusing for customers if company performance rewards/penalties are contingent on a separate gate mechanism. It also removes an element of management flexibility to manage performance commitments across the organisation, and could also drive the wrong behaviours once the gateway has been failed, i.e. if a 'gate' was unlikely to be achieved, there would be a significantly weakened incentive for the company to try to recover performance in that area or any linked area.

We recognise that companies can use various means to ensure that performance commitments are stretching. This is a complex area, and it is important that such tools are used appropriately, i.e. that they are understandable by stakeholders, they truly help to inform customers about company performance and they provide appropriate incentives on companies to improve performance.

We also note that how all such mechanisms are applied, and how stretching performance targets are, all contributes to the overall balance of risk and reward to the company.

Copyright © United Utilities Water Limited 2016

Q2: What is your view on the common performance commitments we are suggesting for PR19?

We support the principle of common performance commitments, where it is appropriate and feasible to do so in a way that provides:

- meaningful information about company performance to customers;
- a helpful basis of comparative performance to CCGs, in holding the company to account; and,
- appropriate, and fair, incentives to companies to improve performance.

For this to be the case, it is important that common measures are consistently measured (i.e. that reported performance is truly comparable and not reflecting differences in measurement methodology) and that all companies can reasonably be expected to have equal opportunities to achieve common performance levels – i.e. there are no significant external environmental factors which mean that companies cannot achieve similar performance for a comparable level of cost.

The application of common performance levels also assumes that each company has historically recovered from customers sufficient revenues to finance the delivery of a comparable levels of service performance and that their inherited networks and assets were equally effective whereas this may not have been the case. Previous serviceability targets were based on companies achieving stable performance against their own historic performance, not the performance levels of another company.

Use of common measures also needs to be considered in the context of support from customers – in particular whether customers in general support common measures and/or targets, and whether these common measures best serve the interests of customer legitimacy.

New customer experience measure

We support the development of a new customer experience measure – our views on Ofwat's thoughts on how this could be developed are set out in response to Question 6.

Water quality compliance

This is a basic hygiene factor for customers – there is no reasonable target other than 100% and, given the industry performance is generally > 99.5% it does not appear appropriate to differentiate companies based on very marginal differences in performance. Also given that it is the DWI's role to monitor and enforce water quality compliance, it may not be appropriate to duplicate this by linking it to a reward/penalty, and perhaps should be reputational only. Furthermore, the definition of this performance measure would need to align with the new DWI proposed measure - however it is as yet unclear how the new measure will be constituted, and it may not be appropriate as an ODI. Our current view of the new DWI measure suggests that it will have a level of subjectivity which is likely to limit its useful application as a comparative measure with a common target.

Customer experience of water quality occurs through the consequential impact on service – e.g. through boil water notices, interruptions to supply, etc. Therefore, we propose that a more customer focused approach to a water quality outcome would be to measure company performance in the occasional event that there is a service outage resulting from compliance issue, e.g. as part of the measure on customer supply interruptions.

Customer water supply interruptions

We support that this should be a common performance measure – loss of service is a key measure of delivery of the water service to customers, regardless of the cause of the outage.

Work done on convergence measures has demonstrated that there are variations in approach to this measure between companies which would need to be resolved if we are to have a true common measure – however, we believe that we are currently close to a commonly defined approach.

Copyright © United Utilities Water Limited 2016

Water distribution input

We support the use of a common water efficiency outcome, and support the use of distribution input.

We recognise the importance of having a strong measure to incentivise reductions in leakage, but the work on convergent measures has revealed some significant issues with the use of leakage as a common measure due to differences in measurement methodology which are related to the size of the company, its supply area characteristics e.g. connection density, topography, deployment of different leakage management packages. Much more work is required to assess whether there is a single methodology that would be appropriate and would allow to account for all these differences in a fair and transparent way. We therefore believe that distribution input is a better approach for a common measure than leakage (or per capita consumption), and for companies to have a bespoke measure and performance level for leakage to reflect their own measurement methodologies and economic performance levels.

Abstraction incentive mechanism

We fully support this as a common performance measure.

Customer property sewer flooding

We support this being a common measure, but do not support there being a common performance target for all companies.

The work on convergence measures has demonstrated that there is still much variability in reporting of sewer flooding performance across the industry. We believe that some external horizontal review is required to ensure that measurement is comparable. Some variations are simply due to methodology differences and therefore common targets wouldn't necessarily drive the right behaviours. This is significantly impacted by regionally variable factors such weather, proportion of densely populated areas, soil type, and sewer misuse – ideally, sewer flooding targets should be normalised at a sustainable "economic" level, analogous to leakage.

Wastewater pollution incidents

In principle we support this as a common performance measure. Any associated ODI should avoid duplication of penalty mechanisms with the EA who will also take enforcement action enforce pollution incidents. We recognise the additional role of ODIs in customer protection (over and above the enforcement role of the EA), but any potential penalties resulting from under-performance should not result in a form of double jeopardy for the company.

This measure looks to combine category 1, 2 and 3 incidents – we support this, as categorisation depends on the sensitivity of the receiving waters which may vary between companies/regions.

The definition should also consider the potential impact of external factors that affect services beyond those that it is reasonable for companies to be fully resilient to, such as third party actions, severe weather or other exceptional events that we identify. For example, it is unclear whether events identified by increased network monitoring, which might otherwise have gone unnoticed and hence unreported, will result in consistent reporting overall for each company.

We also believe that greater weighting should be placed on data reported by companies in the event that national level reports reflect regional variations in reporting. We note that the majority of reported pollution incidents from across the sector are self-reported by the company (e.g. 69% of pollution events in 2015 were reported by the companies themselves), which is also tracked by the EA. It should not be the case that a company could appear to improve performance simply by not reporting them.

Copyright © United Utilities Water Limited 2016

Asset health measures: Water – pipe bursts, and Wastewater – sewer collapses

In principle, "asset health" measures can provide a helpful indicator to customers of how effectively companies have historically invested in the maintenance of infrastructure that provides water and wastewater services to customers. However, the use of a single simplistic measure (such as pipe bursts and sewer collapses) for this purpose (a) implies that the overall health of water network assets is only indicated by this one measure, when in reality it is much more complex, (b) risks inappropriately incentivising companies to focus on that measure alone, to the exclusion of other interventions to improve overall asset health, and (c) is a historical (backward) assessment of performance rather than utilising forward looking leading measures.

The principle of an asset health measure is analogous to the historic use of "serviceability" measures whereby companies were incentivised to maintain a stable level of service, as indicated by a basket of different performance measures. This was also tested by reference to historic company performance. It would not be appropriate to set a common UQ target immediately, without providing a suitable glidepath to achieve this and without understanding the cost of achieving the required change in performance. Absent such information it is also unclear whether that change in performance (and cost) is something that customers actually want to pay for.

In relation to the specific measures proposed:

- **Water** - the impact of mains bursts is already captured within the more customer focussed interruptions to supply measure and so represents duplication of reporting. It should also be recognised that there might be a contrary incentives stemming from other measures – for example if we are incentivised to further reduce leakage then we are likely to find more pipe bursts (as a result of increased activity to find and repair leaks.) Working to reduce leakage would therefore lead to an observed deterioration in performance against the pipe burst measure.
- **Wastewater** - We recommend sewer blockages as a better measure than collapses, as this is the most common cause of service disruption for customers. If Ofwat decides to use sewer collapses, then the measure should exclude "proactive" collapse repairs, which represent a proactive approach to sewer maintenance by companies and should not be included as a service failure. It is also important to be clear on the definition of a sewer to include or exclude rising mains. We also note that use of a sewer collapses measure implies some duplication with the pollution incidents measure and believe that there may be merit in using blockages, rather than collapses, at the measure of performance.

It is also important to recognise that asset health measures are inherently backwards looking, and don't necessarily provide a reliable indicator of the company's current and recent performance in maintaining asset health.

New measure(s) of resilience

In principle we support a common measure on resilience. We set out our views on how this could work in more detail in response to Question 8, and in Annex 1.

In summary, we believe that a simplified, commonly defined risk assessment could enable companies to assess a weighted risk score such as "proportion of customers at high resilience risk". This would enable companies to set a baseline position and compare that to a service level for resilience risk. Improvements in this "population at risk" score can then be used in CBA assessments as part of assessing company proposals on improving the resilience of service in their area. This would constitute a first step towards assessing and measuring resilience on a consistent basis, and therefore one would not expect companies to have comparable performance levels, no matter how consistently the risk assessments had been performed (although it seems likely that any such risk assessment will entail a certain level of subjectivity). Although a horizontal audit could potentially increase the degree of comparability, the trade-off between cost and improvements in resilience are likely to be very company and/or regional specific.

We believe it is important that any comparative measure is risk based as any other type of resilience measure based on current performance or a limited component of resilience could drive the wrong

Copyright © United Utilities Water Limited 2016

behaviours. Only a risk based measure would enable resilience risk to be managed effectively for long term benefit at least cost to customers. With a risk assessment the challenge is attaining consistency in the measure, our suggested approach set out in Annex 1, suggests a series of questions that generally require a "yes or no" answer, which most companies should readily be able to complete from existing data. Whilst there is some work to answer these question for all key assets, once complete it will provide a comprehensive data set from which to assess and plan future priorities.

By undertaking the assessment for multiple hazards, across multiple sites, this will enable companies to understand which hazard(s) represent their greatest driver(s) of risk and which asset groupings are most vulnerable. It will also enable companies to assess whether that risk is spread equally across their region or concentrated on specific populations. By taking a risk based approach, a cost benefit assessment can be undertaken for all intervention types, enabling redundancy, resistance, reliability and response and recovery to be considered on an equal and consistent basis.

Copyright © United Utilities Water Limited 2016

Q3: What is your view on how we might apply comparative assessments at PR19?

In the consultation, Ofwat raised seven aspects of assessing common measures. Many of these issues are interrelated. The most important consideration when making decisions over setting benchmark performance levels is how effective they will be in incentivising efficient performance by all companies. It should be expected that rewards should be (within reason) equally achievable by all companies – i.e. any bias that might prevent a company from achieving the benchmark should be removed where practicable.

The incentive properties of the assessments also should avoid companies making short term investment or operational decisions in the interests of "chasing the measure" rather than incentivising sustainable, efficient improvements in performance across all areas of service.

We consider each of Ofwat's seven aspects in turn below.

1) What type of assessment should be carried out (upper quartile, frontier etc.)

Company performance will naturally be variable, for many reasons in addition to differences in performance, such as:

- differences in measurement and methodology;
- the impact of external conditions in the year (e.g. differences in weather);
- permanent environmental differences which impact on performance (e.g. level of urban or rural populations, industry, source types etc.); and
- differences in historic asset inheritance and investment.

These natural variations result in greater variations in reported performance - i.e. the overall spread between "best" and "worst" performance.

This additional non-performance related variation in any measure of performance means that any percentile, or "frontier" related target is likely to overstate that level of performance relative to its true value – i.e. the observed upper-quartile value of performance is certain to be more stretching (in general) than the true upper quartile measure of performance. More importantly, the variation will manifest differently for different companies, with some finding the upper quartile target easier to achieve (if their variations result in an improvement in performance) and others finding the target impossible to achieve. For this reason, it is imperative that measurement errors and inconsistencies are (as far as possible) eliminated via the use of standardised measurement methodologies, assurance (including horizontal assurance) and, where necessary by:

- setting economic based performance targets – reflecting that the cost of meeting comparable levels of performance is very different in different regions (e.g. leakage, and sewer flooding); and
- allowing exclusions for issues that impact on performance where the source, and mitigation, is outside of management control.

We consider that Ofwat should not go beyond upper quartile targets for PR19. Rather that it should work with the industry to improve the consistency, reliability, and comparability of performance measures over the next five years and revisit this issue at the next price control.

It is also important to consider the interaction between the approach to setting targets for cost assessment and the approach to assessing committed performance levels. In many areas, improvements in service are largely only possible with increased levels of investment (e.g. sewer flooding.) Therefore a true upper quartile performance level should represent the company with upper quartile aggregate performance against all areas – i.e. against a basket of measures and also against cost. Setting all targets independently will likely represent a level of both cost and service performance that no single company could be capable of achieving..

There is also an important interaction between the impact of a change to the benchmark - and the expectation value of rewards and penalties - and overall expected returns. Moving from upper quartile to (say) frontier, would (everything else remaining equal) result in lower overall expected returns, which will

Copyright © United Utilities Water Limited 2016

impact on the overall acceptability of determinations. This would seem to result in a greater weight being placed on returns from financing (via WACC) as the source of company returns, rather than from service performance.

2) What performance information should the assessment be based on: historical, forecast or dynamic?

We do not support the use of dynamic targets, as this would likely to result in inefficient outcomes for customers.

On the one hand, we accept that performance targets that are set dynamically, based on year on year performance would likely provide a more stretching performance target. However, it would almost certainly result in some companies being penalised for good, efficiently delivered levels of service performance in line with customer expectations. It would also result in companies inefficiently investing and prioritising management effort on chasing a moving target, to the detriment of overall service delivery to customers. In particular, a "dynamic frontier" target could result in all but one company in each year being penalised, regardless of whether that level of service was reasonably and efficiently achievable by all companies, and regardless of whether customers supporting companies investing in meeting that service level.

We consider that historic performance is the only reliable basis for assessing relative performance between companies, but that forecast targets should be used by companies as a means by which they demonstrate that they are setting stretching targets, in the interests of providing better service to customers.

3) Whether commitment levels should be set for annual or multi-year averages

We consider that annual targets should be set in preference to multi-year averages – customers experience the impact of within year performance and it seems far more appropriate that companies should be held to account for the quality of service provided in an individual year.

It might be perceived that multi-year averages allow for more variability in service performance to be smoothed out over time. However, we consider that this fails to hold companies to account for elements of service that are within their control, whilst also not properly recognising issues that are outside of management control, and would better be dealt with via targeted exclusions.

It should be noted, as set out in response to question 1, that annual targets are inherently more challenging than multi-year averages, and this should be recognised when designing measures and setting targets. It becomes more important that common measurement and methodologies are applied by companies to ensure that reported performance is being represented on a truly comparable basis, and further that any external factors effecting the economic level of service performance (such as regional weather patterns) are taken account of when setting targets.

4) Whether assessments should apply to individual measures or a basket of measures

In principle we support the notion of a basket of measures which better provides a more fulsome assessment of performance in delivering services to customers, rather than focusing on a single measure which may inappropriately incentivise companies to inefficiently meet a single target than have regard for service provision more generally. Baskets can also enable companies to be targeted on both leading and lagging measures of performance against a single aspect of service, in a way that better holds the company to account for delivering effective performance.

However, we recognise that such an approach is not without problems. Superficially, a "basket of measures" (as with index measures, which are effectively a form of basket) may appear to be less transparent and less easy for customers to engage with than a simple single performance metric. However, it could be that companies still had individual ODIs but that targets would not be at upper quartile for each measure individually - individual ODI targets would be set locally to achieve overall upper quartile performance against the basket index. This would achieve a balance between nationally-set stretching

Copyright © United Utilities Water Limited 2016

targets and responding to local circumstances. On this basis, the ODIs would not be any more complex than now.

5) Whether common commitment levels should be set for the performance commitments only or also set common penalties and rewards

We recognise that common rewards and penalties could provide simplicity and consistency for customers, and feel more legitimate, as they would enable customers to observe that they are compensated the same as customers in other regions for equivalent service failures.

However, this approach would not take account of regional differences in how services are actually valued by customers, the costs involved in achieving a given service level in a given region and what service priorities should be. It is possible that a company's performance incentives could result in improvements in service (and hence rewards) that were not supported by their customers.

6) Whether there should be a time period before the common commitment applies (a glidepath)

In our view, this should depend on the measure. The key consideration should be whether it is reasonable to expect that companies should be able to achieve common performance levels in a single year without (say) significant investment. Where this could not reasonably be expected – such as in the case of sewer flooding, a glide-path would seem more appropriate.

It is also important to recognise that company performance has historically been incentivised based on delivery of stable performance, not on the basis of a common performance level. In order to now achieve common performance levels, some companies may need time and investment to achieve that common level of performance, potentially across more than one AMP.

7) Whether there should be any adjustments for company-specific factors

We consider that there are considerable company specific factors which affect the relative ability of companies to achieve a common level of performance and/or a common level of cost. First of all, there will be a trade-off between cost and performance level. However, there are also impacts from external environmental factors that affect the efficient balance of service performance and cost, such that in more adverse environmental conditions (such as higher rainfall, impacting on sewer flooding) one would expect the efficient service level to be lower, and at a higher cost.

Copyright © United Utilities Water Limited 2016

Q4: To what extent do you agree with our proposed approach to leakage performance commitments for PR19?

As noted in our response to question 2, we recognise the importance of having a strong measure to incentivise reductions in leakage, but the work on convergent measures has revealed some significant issues with the use of leakage as a common measure due to differences in measurement methodology which are related to the size of the company, its supply area characteristics e.g. connection density, topography, deployment of different leakage management packages. Much more work is required to assess whether there is a single methodology that would be appropriate and would allow to account for all these differences in a fair and transparent way, which would require much more work to resolve. We therefore believe that distribution input is a better approach for a common measure than leakage (or per capita consumption), and for companies to have a bespoke measure and performance level for leakage to reflect their own measurement methodologies and economic performance levels.

Furthermore we propose that the distribution input common measure should be a reputational ODI, with leakage as a company bespoke measure with company specific financial ODI linked to the Water Resources Management Plan. This application of a financial ODI for leakage recognises that customers value leakage – however it needs to be bespoke as it isn't currently possible to calculate leakage in a standardised consistent manner. It is also not appropriate to apply an upper quartile target for leakage as the need for investment in this area is subject to a considerable number of external factors, such as the level of water stress and customers' willingness to support investment in this area. An upper quartile target would also not reflect he economic level of leakage and take account of the historic levels of investment.

Ofwat has also raised a proposal for applying a three year average – we do not support this, and consider annual targets to better reflect company management of leakage.

Ofwat has set out a number of further considerations which we have addressed in turn below.

Requiring companies to explain whether the sustainable economic level of leakage (SELL) they are using when proposing their commitment levels takes account of projected future improvements in leakage reduction efficiency

We support this proposal.

Whether the commitment levels take account of long-term expectations of the future value of water, the scope for water trading and the resilience of supplies

We support this proposal.

Why their leakage commitment level is appropriate in the context of comparative information on leakage per property per day or leakage per km of main per day

We recognise the need to explain to customers (and other stakeholders) the company's proposed leakage commitment levels in the context of comparative information, notwithstanding the inconsistent bases on which companies are reporting leakage.

Copyright © United Utilities Water Limited 2016

Q5: What factors should we take into account in our guidance on setting performance levels for bespoke performance commitments at PR19?

In the consultation, Ofwat raised four issues relating to setting performance levels for bespoke measures. We consider each of these in turn below.

1. **Whether there are particular areas that all companies should cover with their bespoke performance commitments, but without requiring a common performance commitment**

Bespoke measures should be linked to customer engagement, CCG views, investment requirements and have some natural progression from AMP6 (lessons learnt and future development).

Company strategy should drive innovation to deliver what customers want from their water company.

2. **Whether, for such areas, Ofwat should require companies to report against common metrics, in addition to their bespoke performance commitments**

In principle we are happy to support this, as companies should be transparent with their reporting against all appropriate measures. However, caution should be taken to avoid too many duplicative measures which could be confusing and potentially conflict with the bespoke measure.

3. **Whether companies must propose performance commitments and ODIs that relate to each price control, such as water resources and bioresources**

We agree that it would be preferable for performance commitments and ODIs to be price control specific, as that would tend to ensure that the individual price controls are held as independently binding. The main consideration should be which party should have the primary accountability for the outcome, and which price control is best placed to exert management control over performance against that outcome.

In some cases, these accountabilities and responsibilities will cross price control boundaries. In such circumstances, it may be necessary to split the ODI into component parts.

4. **Noting that it can be difficult for customers and CCGs to challenge a company's estimate of its marginal or incremental cost, whether companies should be required to provide more information to customers and CCGs about how they have calculated their marginal or incremental cost**

Whilst this seems superficially helpful for CCGs, it seems unlikely any costs provided would be consistent between companies, and therefore this information could equally be as misleading as it is informing. It is also unclear whether Ofwat expects such information would be provided as part of PR19 business plan submissions, through the Annual Performance Report, or bilaterally between companies and CCGs.

Copyright © United Utilities Water Limited 2016

Q6: What is your view on our development of a new customer experience measure for PR19?

Summary

In 'Appendix 5 - A new customer experience measure' Ofwat has set out the context for a future customer service measure. This can be understood as also representing a series of core objectives for a new customer service measure, including:-

- Driving greater customer focus and service levels;
- Supporting trust and confidence in the industry from customers and stakeholders;
- Ensuring a future measure works well within the wider evolving outcomes framework;
- Supporting customer service innovations and developments, enabled by new technology and digital advances;
- Adapting to changing customer expectations and behaviours.

We recognise and endorse this context and objectives.

In our consultation response below we have set out a number of observations and suggestions for achieving these objectives in the design of a new customer service measure. In particular we believe it is important that a new approach:-

- Reflects how customer experience is influenced by factors beyond simple satisfaction, including considering what role measures like trust and ease of interaction might usefully play.
- Learns vital lessons from the design of SIM, including but not limited to ensuring statistically robust sample sizes are captured, survey designs are consistently applied across all companies, and the need to move beyond phone only survey techniques.
- Seeks to understand how industry performance compares to the very best service from other utilities. We would also encourage Ofwat to go further and seek to understand how the industry compares to the very best service providers across all retail markets in the UK.
- Looks beyond traditional contact channels, and instead considers the full range of current and future digital communication channels.
- Retains some element of qualitative metrics, particularly metrics focussed on customer complaints as opposed to the current focus on 'unwanted' contacts.
- Looks to understand customer experience for key customer sub-groups. Particularly customers in vulnerable circumstances and customers in debt.
- Considers the merits of introducing a metric focussed on levels of awareness of companies' vulnerability support schemes, as we believe such a measure provides an effective and pragmatic way of promoting vulnerability support schemes without requiring centrally mandated definition of vulnerability activity levels.

Considering what the focus and outcome should be

> *We support a focus on satisfaction and trust measures. There are also a range of other metrics that may be worth consideration.*

We agree that looking beyond measures of customer satisfaction could add depth to our understanding of customers' experience. Of the various sub elements of customer experience discussed in the consultation we agree that measures of satisfaction and trust are most recognisably relevant to domestic water customers.

Many companies already collect such information, and there are many examples already in use. For example United Utilities has for a number of years run a tracking exercise which provides a quarterly view of customer perceptions of United Utilities compared to a number of service providers from other sectors (banking, telecommunications etc.).

We've been running brand tracking surveys since 2011 by collecting quantitative data from our consumers on a regular basis. For the last 4 years the survey has been run each quarter. The survey enables the company to measure a range of customer perception factors including satisfaction and attitudes to the

Copyright © United Utilities Water Limited 2016

brand. Every survey consists of a sample of 1,000 household customers, both metered and non-metered with a representative socioeconomic and geographic mix.

To keep track of our brand health we measure overall customer perceptions concerning whether our customers understand what we do; how satisfied are they with our service; and what they feel about the way we treat them by rating us. We also ask customers what they think about our 'personality' and rank our traits against other major brands.

Example of our brand tracker personality rankings (anonymised relative to comparator organisations a to i)

Rank	Arrogant	Efficient	Friendly	Genuine	Helpful	Impersonal	Misleading	Reliable
1	*g*	United Utilities	*c*	*d*	*c*	*i*	*i*	United Utilities
2	*i*	*c*	*d*	*c*	*d*	*a*	*a*	*d*
3	*a*	*d*	*h*	United Utilities	United Utilities	*f*	*g*	*c*
4	*b*	*h*	United Utilities	*h*	*h*	*b*	*b*	*a*
5	*e*	*a*	*e*	*e*	*i*	*g*	*e*	*h*
6	*h*	*g*	*f*	*b*	*e*	*e*	*h*	*g*
7	*f*	*f*	*g*	*a*	*a*	United Utilities	*f*	*e*
8	United Utilities	*b*	*a*	*f*	*f*	*h*	United Utilities	*b*
9	*c*	*e*	*b*	*g*	*g*	*c*	*c*	*f*
10	*d*	*i*	*i*	*i*	*b*	*d*	*d*	*i*

The survey allows us to track how our customer engagement campaigns are performing and to measure customer sentiment. For example, key traits associated with satisfaction and trust include customer appreciation; good reputation; socially responsible and not being arrogant or misleading.

We are also measuring how concerned our customers are about their household budget being able to cope with water bills, this helps us to understand affordability issues which many of our customers favour.

Our experience with our brand tracker has helped us see that there are other elements of service that may have value. For example 'ease/effort' measures and 'net promoter' measures are both used widely and can be revealing sub factors of customer overall experience.

Measures of customers' ease/effort of interacting with a company can be a revealing measure, especially where customers do not have a choice of supplier. These measures can assess how easy a company is to do business with, and may be particularly relevant as the water industry adapts to the growth in self-service/digital channels. Some customers' preferred channel of engagement is changing, whilst others continue to find they prefer more traditional engagement routes. Measuring customers' ease of interaction could reveal differences in companies' ability to provide all customers with each's preferred engagement channel.

Net promoter measures are also widely used, and may prove a useful proxy for a range of sub-elements of customer experience. In particular, as revealed to us through brand tracking efforts, customers' views on how much they trust a company and how loyal they feel towards it influence their likelihood of promoting a company with friends and family.

Copyright © United Utilities Water Limited 2016

We think having a separate measure for customers in debt could help address some issues with the current satisfaction measure.

We also see value in developing approaches to testing a wider range of customer segments. Ofwat has drawn out the experience of vulnerable customers as an area of special consideration, a move which we support and discuss more fully below. But there may be other groups of customers where segmentation of customer experience measures may add value.

In particular we think having a separate measure for customers in debt could help address one of the more prominent limitations of the current SIM. Currently companies have good financial incentives to try and collect from customers in debt, but they are also subject to a strong customer satisfaction incentive.

Current evidence suggests that frequent and repeated contacts with customers in debt is crucial to increasing propensity to pay, but also increases the number of debt related customer contacts. We have observed that debt related contacts generally generate lower satisfaction scores from customers than other forms of contact, meaning increased debt collection activity will, all else being equal, reduce levels of debt but also reduce a company's SIM score. This means that all water companies must make a trade-off between robustly pursuing outstanding debts on one hand, and seeking to avoid difficult or negative interactions with customers on the other.

An alternative approach is to measure and report customer experience for customers in debt separately from other customer groups. Such an approach would ensure that the customer experience scores for customers in debt can easily be understood in the context of company debt collection activities. By separating measures of satisfaction for customers in debt, but giving the segment an equal weight to all other customers, a number of difficult counter acting incentives on companies can be resolved, whilst still providing protections for customers both in and out of debt.

We agree that measuring domestic customers' views on 'value for money' in the water industry is not desirable.

We share concerns expressed in the consultation about the reliably of measuring 'value for money' in an industry where consumer choice of supplier is not possible. Evidence is beginning to emerge that there may be a substantial difference in water customers' stated and revealed preference valuations for water services. Such a difference can limit the usefulness of a 'value for money' measure based on stated preferences.

For example, following a recent boil water event in our region we provided compensation to customers of £50 - £60 following impacts on supplies that lasted between 22 and 31 days. Feedback from customers indicate that the level of compensation was generally (although not universally) acceptable. This could be taken to indicate customers value the ability to not need to boil their water at around £2.10/day, or a little under £770/yr. Given that the average household charge for water services in UU's region is only £200/yr we might expect customers view the service as offering high 'value for money'. However through our brand tracker we observe only 51% of customers stating they believe the service offers value for money.

The reasons behind such marked differences in revealed and stated preference studies would need to be better understood if the industry was to start using a 'value for money' stated preference measure to incentivise company behaviour.

Designing a measure of customer trust must be done carefully, but is likely to be worthwhile.

Understanding customers' levels of trust towards individual water companies can add value. But when designing such a measure there needs to be a clear focus on what we are trying to achieve by measuring, comparing, and reporting performance in this area.

When developing a trust measure it is important to understand how customers' views can and are influenced by events external to an individual company. Media coverage can be as much an influence on trust as management action, particularly in the short term.

Copyright © United Utilities Water Limited 2016

For example using our brand tracker we were able to observe the impact on trust during the time that a six week TV documentary; Watermen, was aired by the BBC. This documentary followed the working days of several UU employees, providing an insight into the role and services that water and sewerage companies provide.

We saw positive movement in our trust measures during the six weeks that the TV documentary aired. Measured in June 2014, we saw appreciation of United Utilities by those who had seen the programme versus those who hadn't seen more than one episode positively affected by 21 percentage points.

The programme also helped to show our personality, so had a strong impact on our personality ratings around integrity and caring; innovation and service.

This experience demonstrates that companies can in the short term use media and publicity channels to boost trust based measures, but this may not reflect a change in underlying company performance or customer experience. A measure of company trust ideally should be focussed on long term, sustained variance against an industry specific benchmark. In this way a measure can look through the 'noise' of wider national events and trends, and instead focus on individual companies' efforts to promote trust with their customer base.

We agree that designing such a metric will not be straight forward, and it may take a period of trialling and testing to design an effective measure.

We agree that a trust based measure would initially be better suited as a non-financial measure.

Learning from SIM

Looking back at the design of the current customer survey which underpins the qualitative element of the SIM mechanism, and considering some of its perceived limitations can be instructive in designing a new survey and mechanism. In particular the current SIM survey has been criticised for:-

- **Small sample sizes**. Each quarter as few as 180 customer surveys are conducted for each company's SIM survey. Such small sample sets are not perceived by companies as being statistically robust. Over the course of a year the cumulative number of surveys increases to 720, which is likely to yield more robust results. However in practice, company management is seeking to respond to customer feedback more and more quickly, but is impeded from doing so when the quarterly headline customer measure may not be reflective of true customer perceptions. *This could be corrected by increasing the sample size of each survey to an amount which is recognised by all as being statistically robust*.
- **Infrequent surveys**. SIM surveys are only conducted once a quarter. This is felt to be insufficiently frequent to enable management to understand what is driving changes in customer perceptions as typically multiple changes to service offerings will be implemented in any given quarter. *This could be corrected by increasing the frequency of surveys.*
- **Predictability of survey timing**. The current SIM surveys are highly temporarily constrained in terms of when they will be conducted. Companies can predict to some degree when surveys are likely to be conducted. *This could be corrected by increasing the frequency of surveys, and by requiring companies to provide contact records every week (so removing the need for surveyors to 'call' a survey in advance.)*
- **Inconsistent weighting of customer surveys across services.** The impact of each customer survey on each company's final SIM score varies across retail and wholesale services. For WaSCs each retail weighted survey has a greater impact on final SIM scores than a corresponding water or wastewater focussed survey. As a result significant effort must be expended in ensuring each survey is correctly allocated to the appropriate service stream. This would not be needed if each survey had equal weighting in the final company SIM score. *This could be corrected simply by redesigning the SIM calculations to ensure all customer surveys carry equal weighting.*
- **Inconsistent survey design across companies.** Different survey designs are used for different companies. For example the very smallest WoC takes contact records from a longer time period than all other companies, and two other companies pool their customer contact records and survey results meaning their total number of customer surveys is much larger than other

Copyright © United Utilities Water Limited 2016

companies. These inconsistences in survey design should theoretically influence each companies' final results, and as a result may somewhat undermine the reputational incentives of SIM. *This could be corrected by removing as many company specific adjustments as possible. If adjustments cannot be removed then removing impacted companies from the comparative pool would help retain confidence in the survey for all other companies.*

- **Phone only surveys**. The SIM survey only conducts surveys over the phone. Only using phone based surveys will influence the type of customers that are mostly likely to take part in a survey. A well designed customer survey will use a range of different survey channels, including web and face to face survey techniques. *This could be corrected by using a wider range of survey techniques*

Looking at customer service as it relates to different parts of the value chain

We support Ofwat's view that separate measurements for wholesale and retail are desirable and valuable.

We believe that, in the absence of competition, the distinctions between retail and wholesale services and activities are in practice not well understood by domestic customers. In our day to day interactions with customers we see that customers form a judgement about service and experience in the round. They rarely (if ever) differentiate between a 'billing experience' and a 'wholesale service experience'.

However, at the same time we appreciate that the distinction between retail and wholesale helps inform continuous improvement from a management incentives perspective. Currently United Utilities breaks down and reports internally on customer experience by business area.

Ideally a redesigned set of customer experience measures will also be compatible with any planned changes in industry design and structure, including upstream separation and possible domestic retail competition. Therefore there would appear to be merit in either enabling upfront separation of customer experience across the value chain, or at least understanding how such separation could be done in the near future.

We suggest that whilst we do not observe customers holding a clear distinction between wholesale and retail services that this is not necessarily a barrier to separated incentives. Companies in most instances can and do allocate customer feedback to different elements of the value chain. Therefore if wholesale and retail measurements carry the same weighting in any financial or reputational incentive mechanism it would appear fairly straightforward to ask companies to allocate customer experience measures to elements of the value chain themselves, perhaps with some regulatory review and challenge if appropriate.

Comparing service levels in water with other industries

We fully support comparing company service levels with a wider cross sector pool of companies. Whilst cross comparison within the water industry has been effective in driving up average customer service standards we do recognise that there has been a reduction in the rate of improvement of the very best companies. Pushing the industry to deliver great service relative to a wider pool of sectors could help overcome this current trend.

As domestic water customers do not have a choice of service provider they are most likely to compare their service experience from their water provider with the service they receive from other sectors. It therefore makes sense that companies and regulators are focussed on the same comparisons that our customers are making. We agree the most obviously meaningful comparisons are likely to be against network utilities, wider continuous services (such as broadband), and/or other monopolistic services such as rail.

We believe that initially this comparison may be most easily implemented as a reputation only element of a future mechanism. There are some technical challenges in developing robust, statistically significant measures across multiple sectors, especially when the service provided by those sectors do differ in some important ways. We are also aware that current cross industry comparisons appear to indicate that water companies compare relatively well against average performance in energy, telecoms, and rail sectors, and it may be difficult to explain why a new mechanism should offer financial rewards to water companies for delivering a level of service that customers are already receiving.

Copyright © United Utilities Water Limited 2016

But as an industry we should also consider benchmarking even more widely, including against those companies that regularly deliver the very best in terms of customer service. By testing ourselves against those brands who are well established as providing excellent service we can identify the gap not only at an individual company level but how far the water sector is from companies currently considered to be exemplars. Understanding this gap is important as we should as a sector be aware that customers' expectations will be driven by the very best customer experiences, not just the very best utility providers.

Surveys such as the UKCSI report[1] provide an insight into cross sector relative performance of individual sectors and companies.

Asking customers other than those who have contacted or complained to companies

We agree that engaging with the widest possible customer base is desirable as it helps understand the viewpoint of customers who do not regularly directly contact their water provider.

The technical design of any future customer surveys will need careful consideration. Ensuring that the survey design is effective in engaging a wide range of customer groups and is statistically robust, whilst also ensuring the mechanism's incentives on company management are timely and appropriate will be challenging (see observations on 'learning from SIM' above).

In addition, it is important to consider the incentives placed on management by different survey designs. In particular there would appear to be a big difference in incentives resulting from a survey that potentially contacts any customer and one which potentially surveys any individual in a company's region.

If any individual in a company's region may be surveyed then companies will probably seek to engage with all individuals in their region. The routes to achieving this level of engagement would appear limited, and likely include mass advertising campaigns and other generalised engagement activity.

Conversely if only current customers are surveyed the engagement techniques open to management feel more likely to be wider ranging in design and have the potential to be more targeted at individuals, as the best companies understanding of their current customers is far greater than their understanding of the population at large.

An 'all customers' survey could be achieved by entering all customer contacts in year into the survey sample pool, and adding in an additional number of 'no contact' customers into the pool. This could ensure that individuals surveyed are actual company customers, and so incentivise companies to consider more targeted engagement techniques.

Using more channels to identify and gather the voice of the customer

We should support the greater recognition of digital channels in a future service measure. We agree that where a company receives a direct contact from a customer we should be treating all engagement channels equally. Channels such as web chat and social media are growing in popularity year on year and should be recognised in customer engagement metrics.

We think that Ofwat should not be drawn into defining which specific communication channels should be measured, and instead should set out principles of how to record and measure contacts regardless of the channel used. This will help ensure any channels which may emerge in coming years can also contribute equally to customer service measures.

We think there is merit in distinguishing between direct customer engagements and more general social media commentary. Where a customer is seeking to directly contact a company this should be recognised and treated equally across all engagement channels as part of a core customer experience incentive. For those instances where individuals are making more generalised comments on social media and customer details are not available we would suggest that a separate reputational measure be established.

[1] The Institute of Customer Service 'UKCSI: The state of customer satisfaction in the UK' Jan 2017

Copyright © United Utilities Water Limited 2016

There are well established social media measurements which track metrics such as volume and sentiment. These 'public' contacts are important but different to direct contacts and merit a slightly different treatment. For example during a recent boil water incident in the UU region there were comments on social media posted by individuals whom appear to be physically located outside of the UU region. Whilst such comments should not be ignored it also seems equally true that they should not contribute to UU's **customer** experience measure as the individuals involved are unlikely to have been customers.

Considering the way in which complaints data might be used within a customer service incentive

We think that we should keep quantitative measures as part of an overall customer service measure. Current quantitative measures provide comprehensive coverage of the customer base over time, whereas qualitative surveys can only ever provide a view from a sample of customer over a snapshot in time.

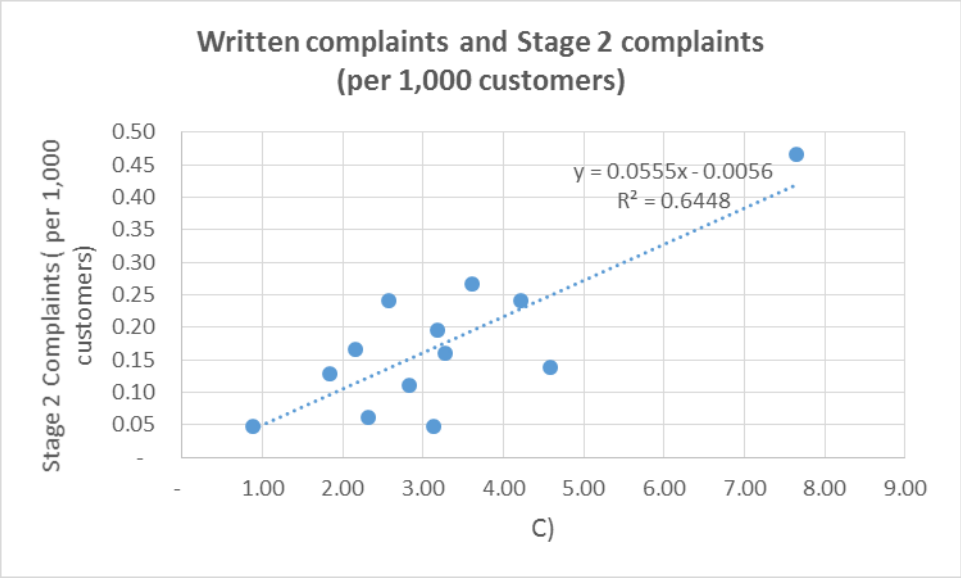
We do agree that the current quantitative SIM measures can in some instances encourage companies to limit customer engagement opportunities in order to discourage unwanted contacts. We recognise that reducing the ease with which a customer can engage with a company can reduce the volume of unwanted contacts received by a company.

We also believe that the concept of an 'unwanted' contact can be difficult to communicate to industry stakeholders, and may well not translate well when considering digital engagement channels such as Twitter and Facebook. We believe there are some indicators where companies also have different (but still compliant) interpretations as to what qualifies as an unwanted contact, creating challenges when trying to fairly compare companies' performance. As a result of these issues we believe that on balance there is merit in discontinuing the collection and reporting of data on 'unwanted' contacts.

In contrast we believe 'complaints' are relatively consistently measured across the water industry. Most stakeholders and customers intuitively know what a compliant is and other industries appear to carry broadly similar definitions of what constitutes a customer complaint (potentially aiding cross industry comparisons). Crucially we believe that it would be difficult to artificially depress the number of complaints received by limiting communication channels with customers. In fact our experience suggests companies which reduce opportunities for customer contact drive up overall complaint levels.

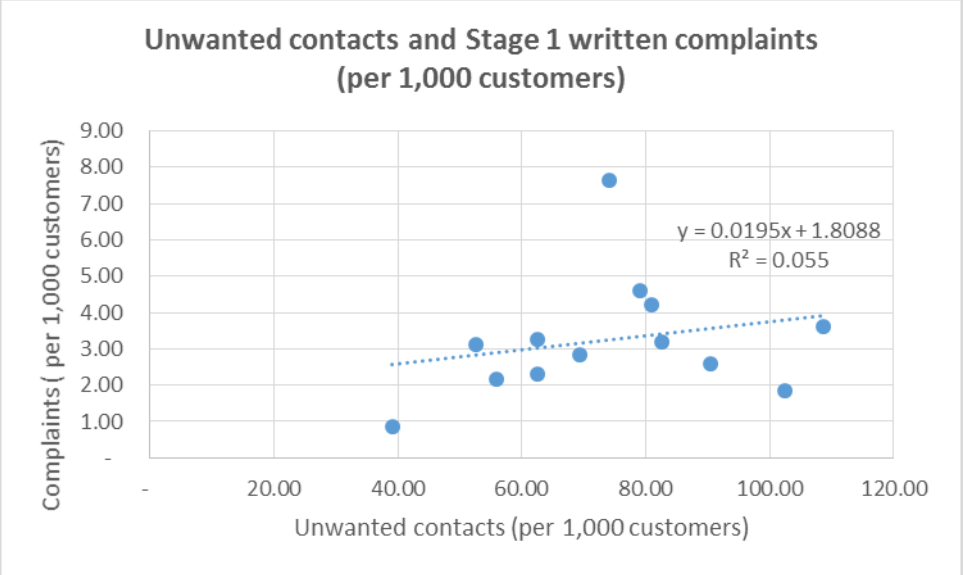
Additionally complaints are indicative of underlying performance to some degree, and occur across the industry at a reasonably high volume. Based on data shared by 13 water companies (with 85% of the domestic customer base) we known that in 2015/16 there were over 80,000 written complaints. This represents a substantial quantitative dataset that should allow meaningful differences in companies' performance to be observed.

Using the same 13 company dataset we can see a good relationship between the number of written complaints a company receives and the number of escalated stage 2 complaints.



Copyright © United Utilities Water Limited 2016

In contrast there is little correlation between unwanted contacts and numbers of complaints, suggesting that unwanted contacts may not be a strong measure of customers' experience.



On balance we believe that a quantitative measure focussed on complaints and escalated complaints can make a substantial contribution to a measure of core levels of customer satisfaction. We see less merit in continuing the measurement of unwanted contacts.

Considering measure(s) to understand engagement with customer in vulnerable circumstances

We endorse the need to consider measure(s)of companies' engagement with customers that find themselves in vulnerable circumstances. Measuring companies' activities in this area is difficult, but we believe it is important for the industry to increase our focus on the support provided to customers in vulnerable circumstances, and regulatory measures have in the past been successful in focussing company efforts in new areas.

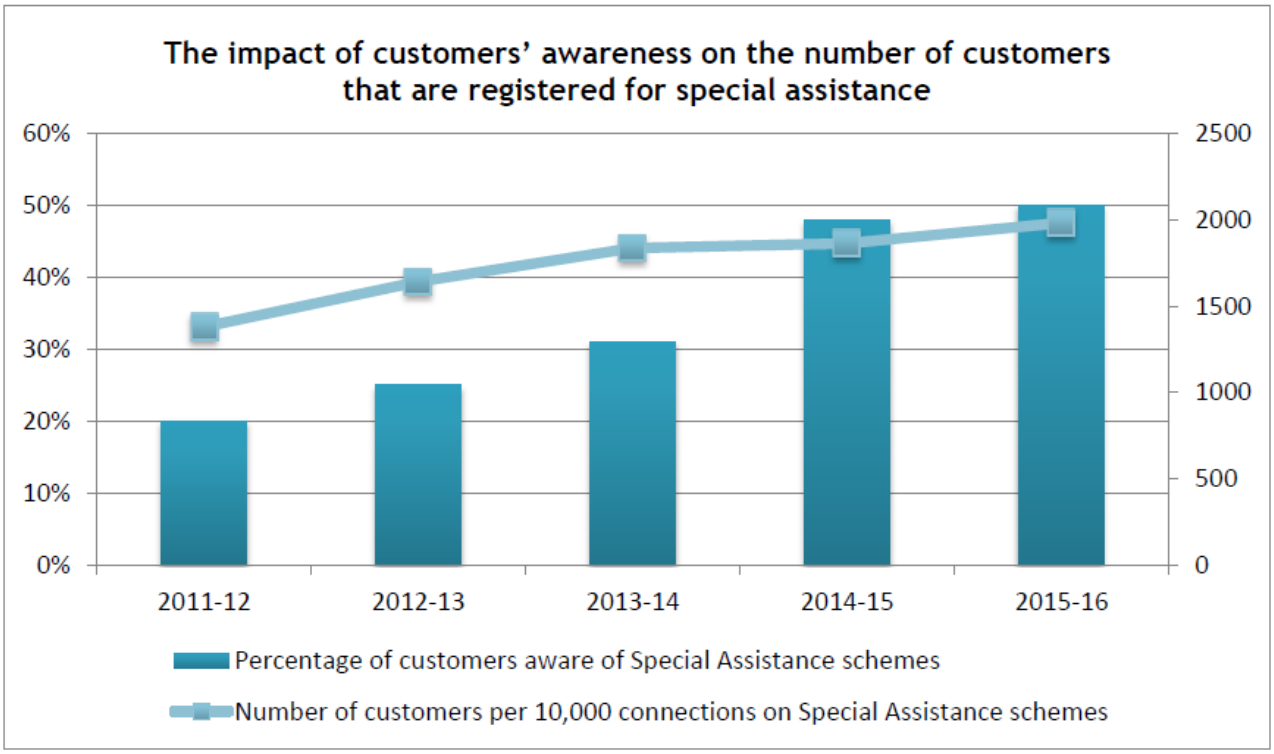
We see particular merit in measures which avoid the need for direct classification of individual customers as 'vulnerable' or 'not vulnerable'. We are also cautious about measures which incentivise company outputs, without focussing on actual outcomes experienced by customers.

Promoting awareness of support schemes

We do see merit in measures which focus on deficiencies in current industry support offerings. In particular we note that one of the most common criticisms of the industry is that too few customers are aware of the support that companies have on offer. In their recent 'Delving into Water 2016'[2] CCWater note that only 8% of customers are aware of the support their local company offers to customers through schemes such as WaterSure. This low level of general customer awareness is a challenge as we know that one common route for customers to sign up for support schemes is via word of mouth from friends and family.

[2] CCWater 'Delving into Water 2016:Performance of the water companies in England and Wales 2011-12 to 2015-16' Nov 2016

Copyright © United Utilities Water Limited 2016



From CCWater 'Delving into Water 2016'

We therefore believe that a metric focussed on all customers' awareness of companies' support schemes could be particularly relevant to the industry at this time. Such a metric would avoid some of the problems with 'output' based metrics (as discussed below), and also avoids the need for the classification of customers into 'vulnerable' and 'not vulnerable' categories.

Focussing on outcomes for customers

As Ofwat set out in its recent vulnerability report[3] customers experience a wide range of different circumstances that will put them into a vulnerable situation. The nature of the circumstances each customer faces can materially alter the appropriate support that they may need at any one time. It is therefore important that any measure of companies' interactions with these customers does not override the need for companies to consider the wide range of difficult circumstances customers face, and the need to ensure the quality of outcomes is at the heart of their interactions with those customers.

Given this we are cautious about the use of output based measures in this area. We do not believe measures such as 'number of customers signed up to support schemes', or 'number of customers signed up for priority service' will drive the right type of company behaviour as there is no focus on the quality of outcomes within these measures.

We are also cautious about asking customer to self-identify as 'vulnerable' as part of a pre-existing satisfaction survey. Many vulnerable customers may be cautious about self-labelling, and many groups of vulnerable customers would likely be poorly reflected in a survey unless the surveyors actively sought out such groups. For this reason, we prefer measures that identify awareness of the schemes available among all customers, as set out above.

[3] Ofwat 'Vulnerability focus report' Feb 2016

Copyright © United Utilities Water Limited 2016

Measures for wholesale only customers

Developer focussed measures

We agree there is merit in considering a customer service measure focussed on developers' experience of wholesale services. Quantitative type measures, similar to the current voluntary scheme already operated via WaterUK would appear a sensible starting position in developing a new measure. These quantitative metrics are established and well understood, and there would appear merit in maintaining some continuity in measures if possible. The industry is investing in improving the comparability of these metrics by commissioning a horizontal audit of company reporting, indicating that companies see value in them.

Requirements for cost reflective developer charges would appear to make linking a financial incentive to the measure difficult, suggesting that a reputational incentive would be most appropriate initially.

Non-household retailer measures

Whilst we recognise that the quality of wholesalers' service provision to non-household customers and retailers is important we believe there are good reasons to wait before developing an additional regulatory measure and incentive.

As set out in the original impact assessment for non-household market opening[4], there are good reasons to believe the newly opening retail market should enable non-household retailers to place coordinated pressure on wholesale services on behalf of their customers without the need for regulatory intervention.

MOSL will be monitoring the operation of the market from its opening. Market and operational performance standards reports will be published which will give objective performance measures and company comparisons. In addition Ofwat's current consultation on monitoring the business retail market from April 2017[5] is also considering similar issues. At this stage we believe the industry should give the market an opportunity to develop before intervening with new regulatory mechanisms.

[4] Defra 1347 'Impact Assessment – Upstream Competition', May 2013

[5] Ofwat 'Monitoring the business retail market from April 2017 - a consultation' Jan 2017

Copyright © United Utilities Water Limited 2016

More powerful outcome delivery incentives

Q7: What is your view on the options for increasing the power of reputational and financial ODIs at PR19?

Consider ways in which the reputational impact of ODIs could be enhanced

In general, the strength of an ODI should be proportionate to the expectation of an efficient response by companies to a change in performance. It should be recognised that simply aiming to increase the power of incentives could incentivise inefficient responses by companies which are not supported by the value that customers place on their water and wastewater services.

In some cases - perhaps more so for many reputational incentives - it may be perceived that there is sufficient visibility and comparability of company performance for companies to feel compelled to act to correct performance. We would consider that the CCGs as well as both the APR and the price control "risk based review" processes provide plenty of opportunity for companies to be held to account for performance against their reputational incentives.

We have also used reputational ODI's historically when testing a new innovative approach or where there is potential overlap with other measures. We therefore seek to ensure that any changes in approach to reputational measures does not stifle opportunities to innovate in developing new performance measures.

Allowing in-period ODIs for all companies which will bring the rewards and penalties closer in time to the service performance that generated them.

Whilst we support the move towards more in-period adjustments, we would note that they might lead to increased levels of bill volatility, particularly if ODIs are being made more powerful. Therefore we believe that companies should retain some discretion to mitigate bill impacts by having some flexibility in the timing with which rewards and penalties are reflected in customer bills, noting customer preferences for bill stability.

Whether end-of-period ODIs (those reconciled at price reviews) should be linked to revenue rather than adjustments to the Regulatory Capital Value (RCV) which take longer to have an impact on companies' revenues.

In principle we do not object to this position, although it would be appropriate to allow companies some discretion about when such ODIs are applied, in the interests of avoiding bill volatility and reflecting customer preferences for bill stability.

Removing the aggregate cap and collar on ODIs which limit ODI rewards and penalties to two percentage points of return on regulated equity.

It is not clear that this is the best way to protect customers from more extreme variations in performance than those predicted at the previous price control. In the event that ODI rewards/penalties exceed a given cap, it could be judged that either (a) the targets set at the price control were not reasonable (either too tough or too lenient), (b) external circumstances caused severe changes to service levels that were outside of the control of the company, or (c) that the company had serious improvement/failure in performance. Whilst in practice it may be a mixture of the above causes, we believe that it is reasonable to set a "soft cap" on rewards and penalties before Ofwat should investigate such changes in performance to ascertain the reasons, and ensure that customers are appropriately protected.

At PR14 a cap of 2% was set which, for UUW, translated to just over 5% on customer bills. This is a material amount and we therefore do not believe that significant increase to the cap is warranted.

Copyright © United Utilities Water Limited 2016

Encouraging efficient companies to hit an overall range for ODI rewards and penalties that was higher than the ranges companies proposed and agreed to at PR14.

Subject to the points raised above, we do not object to this proposal.

We note that at PR14 Ofwat expected companies to present their ODI ranges as a P10 to P90 performance range, and that company approaches to this varied greatly. We consider that it would be far simpler, and less open to interpretation, just to present the maximum possible penalty to maximum possible reward.

Alternative approaches to setting ODI rewards and penalties for PR19 drawing on a wider set of information on customer preferences

We support this proposal.

Adopting industry-standard ODIs for the common performance commitments with powerful rewards and penalties.

We do not disagree with this proposal this in principle, however it is important to recognise that customers in different regions will doubtless have different priorities and therefore different valuations for different aspects of service (including those covered by the common performance commitments). Therefore it is also entirely reasonable that ODIs could differ in strength between different companies, in different regions – it may also be necessary to ensure that incentives on all companies are appropriate. For example an industry standard ODI might incentivise service improvement (at customers' expense) in a region where customers are already satisfied, and/or may under-incentivise service improvement in another region where customer support is far stronger.

Encouraging companies to increase the proportion of ODIs with financial rewards

We consider that what is most important when choosing financial or reputational ODIs is the behaviour that it drives, and whether converting an ODI from reputational to financial is helpful or unhelpful overall (e.g. if it diverts attention away from other areas of service performance).

It is also essential that companies are able to demonstrate that there is sufficient evidence of customer support for a financial ODI, otherwise it risks a misalignment of customer needs and company incentives. That said, it was notable that there was considerable variation between companies on the number of reputational ODIs at PR14. Where similar service areas are being treated differently then it would seem appropriate to challenge companies as to why this is appropriate.

The detailed design of ODIs such as the use of deadbands

It is important to carefully consider the role of a deadband within ODI design in order to ascertain whether or not a particular deadband is appropriate. Deadbands can be a useful tool to manage natural variation in performance around a particular target. Removal of a deadband will most likely act to increase the effective target that companies would need to meet in order to avoid a penalty (i.e. to ensure that all of the natural variation falls above the target). If this is the intention, then the resulting incentive would mean that the true effective target would not be consistent with the published target. If such an approach were followed then it would be more accurate to describe such a target as a "minimum performance level."

Deadbands can also work in conjunction with other elements of ODI design such as exclusions – the absence of deadband would likely result in more fulsome list of required exclusions, whereas a well designed deadband would mitigate against most general risks to performance which are outside of management control.

Copyright © United Utilities Water Limited 2016

The implications for ODIs of an approach based on a variable cost of equity that is partly based on how stretching a company's performance commitments are

Consistent with our response to Ofwat's consultation on the cost of debt, we do not object to the principle of linking overall revenues to the level of ambition in company plans. However:

- We do not support the use of the "cost of equity" for this purpose (to maintain a common WACC value that is not affected by perception of company plans.) Instead, if an incentive is to be applied then it should be a revenue figure which translates to a variable "return on equity".
- It is important to ensure that any perception of how stretching a company's plan is (including its ODIs) is not simply a result of favourable operating conditions which facilitate higher service levels than other companies – it should be genuinely as a result of the company stretching itself, subject to the constraints of operating within its region. We would be concerned if this differentiation were not taken seriously or considered only superficially as it would lead to unreasonable detriment to companies within a more challenging external environment.

The implication of more powerful ODIs, for example, performance commitments will need to be stretching, clearly defined and closely reflect what their customers want.

We fully support the need to reflect customer priorities in setting all ODIs.

Whether 'gated' ODIs could be used, where rewards on some ODIs are contingent on a company incurring no penalties on other ODIs

This seems, in principle, similar to the "basket of measures" approach to an ODI, where rewards against one element of the basket could be offset by poor performance on another element. We also presume that such an approach could only be meaningful if the two ODIs (the main ODI and the "gate" measure) are somehow related e.g. leakage and security of supply. This can have the effect of making rewards more difficult to achieve if the company has only performed in one area (although conversely companies could also avoid penalties if performance elsewhere is good). Gated ODIs would increase the level of complexity within calculations, and may be confusing for customers if company performance rewards/penalties are contingent on a separate gate mechanism. It also removes an element of management flexibility to manage performance commitments across the organisation, and could also drive the wrong behaviours once the gateway has been failed.

Copyright © United Utilities Water Limited 2016

Better reflecting resilience in outcomes

Q8: What is your view on our proposals for better reflecting resilience within the outcomes framework?

Ofwat sets out the following principles:

- Principle 1 - A better and more integrated understanding of service risks
- Principle 2 - Customer engagement
- Principle 3 - A resilience golden thread and greater transparency
- Principle 4 - Broad consideration of intervention options
- Principle 5 - Cost effective solutions
- Principle 6 - Outcomes and customer focussed approach
- Principle 7 - Board assurance and sign-off

Whilst we fully support these principles as reflecting the right general approach, we have some comments on the detailed application of these principles within the overall price control framework, in particular on the derivation of a sensible and usable resilience metric.

Risk assessment, and the golden thread to a resilience metric

We agree that each company should undertake a detailed risk assessment. However the outcome of this assessment will (and should, to ensure there is a golden thread) influence any interventions required, and any measures used within a resilience ODI. The problem is that resilience is not simple or easy to define, nor is it reasonably and consistently reflected in any current simplistic performance metric.

It is difficult to see how a golden thread can be maintained unless the resilience metric itself results from the risk assessment. Furthermore, the metric should reflect the company's position after consideration of all resilience measures in place (i.e. residual resilience risk), not just a gross position of resilience need – that should be considered in the risk assessment opposite measures that the company already has in place.

The suggested metrics noted in the consultation all have some issues or shortcomings, as set out below:

Suggested metric	Issue / shortcoming
Percentage of water treatment / wastewater treatment works at risk from flooding that have protection in place	Organisations need to consider the most cost effective approach to delivering a resilient service. This proposal only considers building resistance to flooding as a hazard, and does not take account of legitimate and cost effective response and recovery measures in place
Percentage of customers with more than one source of supply	This only considers one aspect of resilience (redundancy) and not the three others (resistance, reliability, and response & recovery), and is therefore a measure of gross risk.
Security of supply under drought conditions	Arguably this is a service level, not a measure of resilience
Reduction in volume or proportion of surface water entering the public sewer system	This indicates an activity which results in a change in service which may be more resilient, but does not measure absolute resilience

Copyright © United Utilities Water Limited 2016

Proportion of total abstraction from sources that pose a risk to the environment	This is measuring an externality, not the resilience of the company's service to manage that externality.
Percentage of households eligible for a social tariff actually receiving a social tariff	This indicates an activity which results in a change in service which may be more resilient, but does not measure absolute resilience
Households served on a tariff or financial incentive scheme that rewards water saving	This indicates an activity which results in a change in service which may be more resilient, but does not measure absolute resilience
Acres of land in better stewardship due to water company catchment management	This indicates an activity which results in a change in service which may be more resilient, but does not measure absolute resilience
Total litres/year discharge of untreated wastewater (including from combined sewer overflows) into the environment, normalised per capita served.	This is measuring a permitted discharge, which is due to rainfall, antecedent conditions and the size, age and type (combined or foul) of asset. It does not reflect the resilience of the wastewater service.
Percentage of catchments with a long-term strategic wastewater plan	This indicates an activity which results in better planning, this approach may lead to more resilient planning, but does not measure absolute resilience itself.

Any risk assessment will necessarily be subjective, but we believe that it could be commonly defined, by setting out elements of the risk assessments undertaken by companies which should be consistent. For example:

- **Commonly specified hazards** - the range of potential hazards against which companies are testing their services could be consistently defined - e.g. to all use the same frequency of storm event;
- **Common scoring / weighting** applied to different elements of the risk assessment; and
- **Consistency testing** of risk assessments – e.g. via horizontal review.

We have set out in Annex 1 of this document a proposed indicative risk assessment process (for the water service) which demonstrates how a standardised resilience risk assessment could be established.

The outcome of the risk assessment would also need to be a common (customer focused) measure. This would take a form such as:

- ***the number of customers whose service could not be restored within X hours from specified hazards (and weighted accordingly on a common basis); or***
- ***the number of customers whose service risk assessment score falls below a specified value.***

The latter of these seems more likely to be readily extractable from the risk assessment process.

We believe that this kind of standardised approach would have the benefit of being customer focussed (e.g. on number of customers with a significant duration of service outage following an event), as well as being fully net of all resilience measures already in place, including response and recovery. This also provides for a common measure of resilience improvement when assessing the merits of different options for investment in resilience improvements, and could also support a common benefits assessment.

We recognise that this approach would need some further development, and further that it will likely require some adaptation to be suitable for all aspects of water and wastewater services.

Copyright © United Utilities Water Limited 2016

Q9: What is your view on the options and our preferred approach to asset health outcomes?

We support Ofwat's general approach to asset health, which is to move towards an increased level of standardisation, but for companies to set out their proposals in business plans.

In principle, "asset health" measures can provide a helpful indicator to customers of how effectively companies have historically invested in the maintenance of infrastructure that provides water and wastewater services to customers. However, the use of a single simplistic measure (such as pipe bursts and sewer collapses) for this purpose (a) implies that the overall health of water network assets is only indicated by this one measure, when in reality it is much more complex, (b) risks inappropriately incentivising companies to focus on that measure alone, to the exclusion of other interventions to improve overall asset health, and (c) is a historical (backward) assessment of performance rather than utilising forward looking leading measures.

Experience from PR14 suggests that a greater level of definition is required when setting measures, getting clarity on how each measure will work and any calculation that sits behind it. However we need to be clear on how sub-measures work within target setting - our view is that it is the overall measure that is the target not the sub-measures within a basket. This allows for management balancing of factors in order to maintain the high level target.

We support some increase in standardisation of the approach, but there should be some recognition that assets differ between companies and therefore standard targets may not be appropriate.

Copyright © United Utilities Water Limited 2016

Making performance commitments more transparent

Q10: To what extent do you agree with our proposals for making performance commitments more transparent for customers?

Transparency principles

Ofwat sets out four principles of transparency, which we fully support:

- **Clear** – use plain English as much as possible, avoid jargon, and use technical terms only if they are well explained.
- **Unambiguous** – definitions should leave no room for doubt about what is being proposed or measured.
- **Complete** – the definitions should describe any material points of relevance which a company might later rely on when reporting its performance.
- **Concise** – the definitions should be as short as reasonably possible to enable customers, CCGs and us to engage effectively with the definitions.

We recognise that it is difficult for companies to be held to account for their performance unless it is relatively straightforward for customers and other stakeholders to understand our performance measures. One of the difficulties with this is balancing simplicity and ease of understanding with the complexity of indicators which might be needed to more fully represent performance against a particular aspect of service (as noted in response to question 9 on asset health measures).

Companies to provide better clarity on sub-measures

We fully support this, and agree that companies should provide better clarity and transparency of all sub-measures within an ODI, particularly where a baskets of measures or a performance index is used.

Company communication of performance information

We fully support this – in the case of UUW, data is already shared with YourVoice (the North West's customer challenge group) and externally via our website both via a comprehensive Annual Performance Report and a customer focussed summary with a "Crystal" mark for plain English.

We note that we need sufficient flexibility to allow innovation of the approach to disseminating information to customers, as this might evolve during AMP7. We would not want to limit the opportunity for innovation in the approach we take to engagement with customers by requiring the full detailed approach to be set out at PR19.

How to deal with scheme specific outcomes

In general we are supportive of Ofwat's approach to scheme specific outcomes. It seems to us that scheme specific outcomes are most likely to be sought where there is a specific service enhancement (e.g. those specified by the National Environment Programme) that results in a number of projects to be delivered, or where there is a particular large project whereby a specific ODI governs its delivery to ensure that customers are protected.

In the case of the environment programme where there may be the need for scheme specific performance commitments, some flexibility will be required due to the timing in publication of environmental requirements.

Copyright © United Utilities Water Limited 2016

ANNEX 1 – Resilience risk assessment and metric

Within the water industry over the last 18 months or so, there has been a wide debate about potential resilience metrics. They tend to fall to two extremes, a complex risk framework that is all encompassing but would be extremely challenging to setup in a manner to achieve any kind of consistency between companies or simplistic lagging measures of performance that would be quick and easy to implement but of very limited value in understanding the true resilience of companies.

The challenge is to find a metric that sits at a point between these two extremes, simple enough to define to ensure some consistency and sophisticated enough to be of value in providing at least a pseudo comparison of relative resilience. The table below presents four categories of measures which we expect to be proposed to Ofwat in descending order of complexity and time to complete but arguably also descending order of completeness and sophistication.

Metric type	Examples	Strengths	Weaknesses
Resilience Framework – Set a framework over a range of resilience issues such as leadership, strategy, environment, people and assets and compare companies to a scale of best practice. Similar to an OPA type approach, this could be externally audited and companies compared.	• Arup Resilient Cities framework developed with the Rockefeller Foundation	• Broad assessment of resilience to hazards and trends • Not limited to an asset focus but demonstrates an awareness of wider stresses and trends such as skills shortages and finance • Is not a "capex justifying" measure as it drives policy and procedure change, you can't just build your way to a high score • Allows comparison between companies	• Would need some time to setup the framework and scoring mechanisms • Even with a definition around the rules, it would still be somewhat subjective to score and would require auditing to give confidence in use as a comparator • A more detailed, quantitative risk assessment would still be required to justify totex requirements for assets • Arguable distracts from the highest priority which is risk of service failure, could be a better long term objective for a measure

Copyright © United Utilities Water Limited 2016

Metric type	Examples	Strengths	Weaknesses
Risk Based Assessment – Set a methodology for assessing service risk on an asset by asset basis, this would include consequence assessments and ideally probability assessments either qualitative or quantitative which could be summed to a company level.	• An assessment based on site by site analysis of consequence of failure and vulnerability to hazards, see example below	• Consistent with the common framework approach adopted for asset management planning • Allows comparison at a site by site and company by company level • Can be used for investment prioritisation by allowing cost-benefit assessment • Can demonstrate risk reduction over time	• Complex to setup in a way that would allow comparisons between companies • Even with rules, some subjectivity would remain • Asset focussed and doesn't pick up broader resilience issues
Pseudo Risk Based Measures – some proposals for assessing risk are limited in their assessment of service risk to customers or have the potential to drive the wrong behaviours as they "reward" one type of intervention more than another that could be equally valid and more cost effective	• Percentage of water treatment / wastewater treatment works at risk from flooding that have protection in place • Percentage of customers with more than one source of supply	• Simpler to assess than a full risk based assessment • Would allow companies to track change over time, so could be used to set aspirational long term resilience targets	• Measures miss aspects of a true risk assessment, either take no account of probability or don't consider consequence of hazards • Could drive the wrong behaviours as only certain intervention types would appear effective • Measures usually depend on local conditions, can't be used to compare
Basket of Outcome Measures – using a selection of existing outcome type measures would give an indication of relative resilience performance	• Security of supply index • Customer minutes lost	• Most already being calculated so straightforward to develop once agreed which measures are in	• Most are lagging measures so variation of hazards will be significant score impact, doesn't necessarily mean resilient or not • Many are dependent on local conditions so can't be used to compare

Copyright © United Utilities Water Limited 2016

Many of the measures that have been suggested are not truly risk based and/or not comprehensive. Many are also either lagging measures of performance or they are too narrowly focussed on a single aspect of resilience. For example, the measure "% customers served by a single supply" only considers whether a water supply system has sufficient redundancy. In remote areas, where populations are more widely dispersed, the distances involved in connecting supplies together may be too great to cost effectively provide a second source. Resilience in these areas may be provided by ensuring sites are resistant to hazards, operate reliably and have a robust response and recovery plan. All valid options should be considered when managing resilience of supplies to customers, but many of these would not register on such a simple limited measure. Not all companies could cost-effectively achieve zero customers served by a single supply, they would individually need to assess a maximum achievable target. The target could not therefore be used to compare companies and the risk is that even if that is the intention, comparisons are made and companies may feel pressured to perform to this target and this could drive sub-optimal investment decisions for customers.

With all the debate around resilience in the water industry, there are no shortage of suggested measures. The risk is that individual companies may feel more favourably towards measures that work better within their individual circumstances - hence there is a dominance of drought resilience measures by the companies in the South East. The DWI want measures around treatment works deemed "too big to fail" whereas the Environment Agency see an opportunity to promote more measures related to environmental improvements and vulnerability to flooding. In all of these cases, they are measures of resilience but they are not broad enough to consider the overall resilience of services being provided to customers. Only a risk assessment can consider all of these aspects and potential achieve the required breadth in assessing resilience overall.

Copyright © United Utilities Water Limited 2016

Setting up a consistent risk assessment approach

Any proactive resilience measure should be risk based, so that progress can be measured in reducing the risk to service irrespective of whether incidents occur

Risk = Consequence x Probability

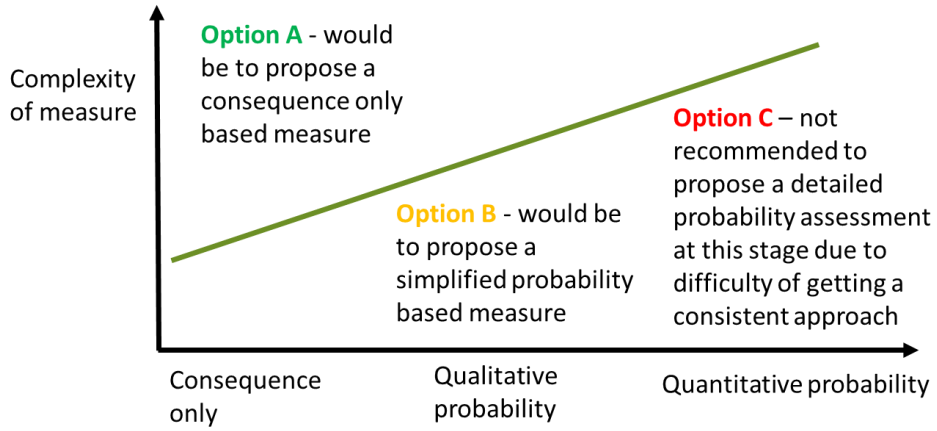
= (Scale of impact x Duration of Impact) x Probability

Probability is the most problematic to set consistently as you have to assess not just the probability of the hazard occurring (based on a fixed severity or return period) but also whether a hazard of that severity would impact the asset. One also needs rules to assess whether or not existing controls are fully effective.

There would be considerable inconsistency and, even with rules, it would be challenging to perform a risk assessment without a degree of subjectivity. A more qualitative assessment might be more realistic but would be less specific and therefore less helpful.

We consider there are three options:

A. **Consequence only assessment**
B. **Simplified risk assessment**
C. **Detailed risk assessment**



We recommend that a simplified risk assessment offers the most appropriate balance between simplicity of standardisation and the usefulness of the data provided in delivering a form of comparison of relative resilience between companies and at least offers a means of tracking change at a company level over time, enabling the cost-benefit of interventions to be assessed and improvement to be seen over time.

Separate risk assessments could be created for water (based on interruptions to wholesome water supplied) and for wastewater (based on flooding of properties from asset failures and pollution from spills and a failure to adequately treat). We have focussed on developing the water measure at this stage so this is presented as a more complete example.

The outline of the measure is that consequence would be assessed as number of properties affected if a site is lost and probability is an assessment of the vulnerability of that site to a variety of hazards.

Copyright © United Utilities Water Limited 2016

Example		Water Treatment Works A		Water Treatment Works B		Water Treatment Works C	
Consequence		Supplies a small town	20k props	Supplies a small city	150k props	Supplies a large city	400k props
Hazards	**Flooding**	High	1	High	1	Low	0.1
	Malicious damage	Medium	0.5	Low	0.1	Medium	0.5
	Power loss	Low	0.1	High	1	Low	0.1
	Fire	Low	0.1	High	1	Low	0.1
	Source loss	Low	0.1	High	1	Low	0.1
	Asset failure	Medium	0.5	High	1	Medium	0.5
	Human effects	Low	0.1	Medium	0.5	Low	0.1
	Telemetry failure	Medium	0.5	Low	0.1	Low	0.1
	Total for site		**3.8**		**5.7**		**1.6**
Risk score		**20k * 3.8**	**76k**	**150k * 5.7**	**855K**	**400k * 1.6**	**640k**

In this example, although Water Treatment Works C serves the larger population it is less vulnerable to hazards and therefore your priority in reducing resilience risk would be to target Water Treatment Works B and particularly those hazards at Water Treatment Works B that are rated as high. This would also enable the cost effectiveness of resilience measures to affect the score and their relative cost-benefit to be assessed. For example:

At Water Treatment Works B, a new trunk main can supply the whole area from another source (redundancy) and the consequence therefore drops to zero but at a cost of £30 million. A permanent flood barrier (resistance) would reduce the flood risk to low at a cost of £1 million. Which offers best value?

- £30m to reduce 855k risk score to zero = £35k per risk point (remaining residual risk is zero)
- £1m to reduce risk score by 135k = £7.4k per risk point (remaining residual risk is 720k)

On this basis the flood barrier offers the best value for money although the residual risk is higher. The analysis is carried out on a site by site basis and then the total risk score is summed to give a company value. The data can then be cut in different ways so we can also assess which hazards represents the greatest threat to service or which asset type is the greatest contributor to risk score, water treatment works or trunk mains.

Copyright © United Utilities Water Limited 2016

Simplified Risk Assessment – Water Service

This approach follows the basic premise that risk is consequence and probability and also assesses the extent to which the 4R's of resilience (as advocated by the Cabinet Office "Keeping the Country Running") namely: resistance, reliability, redundancy and response/recovery influence either the consequence or probability or both.

Stage 1. Define the consequence of failure

Which assets should be included in the assessment?	**Included**	Treatment works (including risk to source and raw water conveyance, could include boreholes, impounding reservoirs[6], intakes and catchments)
		Service reservoirs
		Pumping stations
		Trunk mains (defined as sections of trunk main that share a reasonably common level of risk to service i.e. between offtakes)
		Pipebridges
	Not included	Downstream dosing points, other facilities e.g. laboratories, fleet stores, control centres, etc.
Should the assessment be extended to all assets?		We have applied a *de minimis* exclusion of assets that supply more than 5,000 properties under normal operating conditions on the basis that smaller areas can usually be maintained through tankering

[6] This is just the risk to water supply from loss or contamination of an impoundment source not the risk to populations downstream of an impoundment failure as this is the subject of our "Portfolio Risk Assessment" approach

Copyright © United Utilities Water Limited 2016

Each asset that meets the criteria should be assessed using hydraulic models to see whether, with existing assets, customer supplies can be maintained if the asset is unavailable under a single scenario:

Duration	We have assessed the impact on service if the asset is unavailable for seven consecutive days, at this stage the assumption should be tested that no return to service has been possible.
Demand	Model should use normal demand as this is largely a test of asset failure and the Water Resources Management Plan will test more extreme demand scenarios.
Storage	The model should start with the lower end of normal operating storage typically c.80%.
Rezoning	Where feasible, rezoning should be tested to reduce the consequence of failure (albeit that there may then be uncertainty as to whether valves are functioning and the rezone can be implemented without discoloration risk)

If at the end of stage 1, the asset can be removed and rezoned without impact to customer service then the asset is deemed to be resilient on the basis of existing redundancy and no further assessment of risk is necessary for that site. If customer service is vulnerable to an outage of this duration then the assessment moves on to stage 2.

Copyright © United Utilities Water Limited 2016

Stage 2. Assessment of vulnerability to shocks

The UKWIR "Resilience Planning: Good Practice Guide" lists 46 separate hazards, albeit they are actually a mixture of shocks, stresses and trends. A fully detailed risk assessment would consider all of these and their potential to affect each asset but in a simplified approach we are applying a small subset of these hazards, those deemed most likely to drive risk. We have grouped some of the hazards together as they are similar and have similar controls. UKWIR also acknowledge that this list is not exhaustive and we have added two hazards that are not in UKWIR's list. In our approach we have developed a question set related to eight hazards:

1. Flooding including coastal, fluvial, surface water and groundwater
2. Malicious damage including sabotage, security related and third party intervention
3. Power failure and power loss
4. Fire
5. Raw water loss/contamination including catchments, reservoirs, boreholes, rivers and raw water distribution
6. Asset failure due to condition
7. Human factors including operational error
8. Telemetry failure

More could be added in future if a viable question set can be defined to assess probability of the remaining hazards that may be more challenging to quantify. Each question set has qualitative questions to assess:

- the relative vulnerability or resistance of the asset to that hazard,
- how reliably the asset could continue to perform if impacted by that hazard
- whether there is a response and recovery plan to enable the site to recovery quickly

An example for flooding is presented below, for the first four of the hazards listed above.

Copyright © United Utilities Water Limited 2016

Stage 2a. Assessment of flooding probability

There are a number of potential sources of flooding and each is first assessed separately:

Flood source	Data	Assessment	Score
Fluvial and coastal flooding	EA flood map, Flood Zone 2 (1 in 1000)	Is at least 10% of asset within this flood zone? How many times has the site flooded (from a fluvial or coastal source) in the last 25 years?	If yes, add 0.5 to score Add 0.2 per incident
Surface water flooding	EA surface water flood map (1 in 1000)	Is at least 10% of asset within this flood zone? How many times has the site flooded (from a surface water source) in the last 25 years?	If yes, add 0.3 to score Add 0.2 per incident
Ground water flooding	None	How many times has the site flooded (from a ground water source) in the last 25 years?	Add 0.2 per incident
			Sum, if >1 then =1

Then further questions on the resilience of the site to the hazard:

Resistance	Does the site have permanent flood barrier designed to a 1 in 1000 flood level?	If yes, then remove fluvial, coastal and surface water scores
Reliability	Can the assets continue to function in a 1 in 1000 year flood event e.g. due to raised electricals, etc?	If yes, then reduce risk score to zero
Response and recovery	Is there a site specific plan to recover the functionality (>90%) within 7 days?	If yes, then reduce risk score by 50%

The risk that climate change will increase flood risk is not included at this stage as there is no national data set on which to make the assessment of how flood risk will change over time.

Copyright © United Utilities Water Limited 2016

Stage 2b. Assessment of malicious damage probability

Assessment of vulnerability

Assessment	Score
Is the site urban, rural or semi-rural?	Urban = 0.8 Semi-rural = 0.5 Rural = 0.2
Has the site suffered security breaches in the past 25 years?	Add 0.1 per incident
	Sum, if >1 then =1

Then further questions on the resilience of the site to the hazard:

Resistance	Is the site protected to Critical National Infrastructure "Enhanced" or "Basic" security level e.g. fences, CCTV, etc?	Enhanced - reduce risk score by 80% Basic - reduce risk score by 50% No – no change
Reliability	Is the site staffed 24 hours?	Yes – reduce risk score by 50% No – no change
Response and recovery	Is there a site specific plan to recover the functionality (>90%) within 7 days?	If yes, then reduce risk score by 50%

Copyright © United Utilities Water Limited 2016

Stage 2c. Assessment of power failure and power loss probability

Assessment of vulnerability

Assessment	Score
Are there powered assets whose failure would disrupt supplies to customers?	Yes = 0.6 No = 0
If yes, then:	
Has the site suffered power outages in the past 25 years?	Add 0.1 per incident or assume 1 if unknown
Can the site be supplied from more than one substation?	Reduce by 50%
	Sum, if >1 then =1

Then further questions on the resilience of the site to the hazard:

Resistance	How many independent power supplies to the site?	Divide by number of supplies
Reliability	Is there a standby generator on site?	Yes - reduce risk score by 80% No – no change
Response and recovery	Are there hook up points to connect a mobile generator? Can the site be powered by a mobile generator?	Yes - reduce risk score by 50% No – no change

Copyright © United Utilities Water Limited 2016

Stage 2d. Assessment of fire probability

Assessment of vulnerability

Assessment	Score
Are there any high voltage assets on site?	Yes = 0.5 No = 0
Is there a source of fuel on site e.g. diesel, wood, paper, chemicals, etc?	Yes = 0.5 No = 0
Is there a potential ignition source on site e.g. heaters, lighting, engines, electrics, etc?	Yes = 0.5 No = 0
	Sum, if >1 then =1

Then further questions on the resilience of the site to the hazard:

Resistance		
Reliability		
Response and recovery	Is there an automatic fire detection system linked to a central control?	If yes, reduce risk score by 30%
	Is there a sprinkler system?	If yes, reduce risk score by 50%

Copyright © United Utilities Water Limited 2016

Stage 3: Calculation of site risk score

The "probability" score for each hazard is then multiplied by the consequence to get a risk score for each hazard. Summing the risk scores gives a total risk score for the site. In future, additional hazards could be added to the methodology without needing significant rework. By picking up the hazards most likely to cause a service outage now, other less likely hazards should contribute less to the risk score and therefore adding more hazards improves the accuracy of the measure but is not likely to materially alter the outcome.

Asset risk scores could then be divided by population served to give a weighted risk score. Assets over a defined risk score per population could be defined as high risk.

Stage 4: Calculation of company risk score

Once a risk score per person has been assessed for every asset, the outcome should be represented as customer focused measure. This could take a form such as:

- ***the number of customers whose service could not be restored within X hours from specified hazards (and weighted accordingly on a common basis); or***
- ***the number of customers whose service risk assessment score falls below a specified value.***

The latter of these seems more likely to be readily extractable from the risk assessment process. Therefore the "number (or proportion) of customers served by assets at high risk" would be the (normalised and comparable) resilience outcome measure obtained from the risk assessment.

Interventions could then be prioritised on the basis of cost to reduce the risk score and interventions that reduce the probability or consequence of a hazard leading to a service failure would all reduce the risk score in a consistent manner. This measure would not therefore favour any one type of intervention compared to another and would enable the most cost-effective interventions to be identified and have a demonstrable impact in reducing overall risk.

Conclusion

We believe that this kind of standardised approach would have the benefit of being customer focussed (e.g. on number of customers with a significant duration of service outage following an event), as well as being fully net of all resilience measures already in place, including response and recovery. This also provides for a common measure of resilience improvement when assessing the merits of different options for investment in resilience improvements, and could also support a common benefits assessment.

We recognise that this approach would need some further development, and further that it will likely require some adaptation to be suitable for all aspects of water and wastewater services.

Copyright © United Utilities Water Limited 2016