PR19 final determinations

Customer measure of experience (C-Mex) and developer services measure of experience (D-Mex) policy appendix
PR19 final determinations: Customer measure of experience (C-MeX) and developer services measure of experience (D-MeX) policy appendix
About this document

The customer measure of experience (C-MeX) and developer services measure of experience (D-MeX) are financial and reputational common performance commitments designed to provide customers in the water sector with excellent levels of service.

This document sets out our final decisions for C-MeX and D-MeX as part of our final determinations for the 2019 price review (PR19). C-MeX and D-MeX come into effect on 1 April 2020 and will operate until 31 March 2025. They apply to the 17 largest water companies in England and Wales and are set out in the ‘Outcomes performance commitment appendix’ for each company.

Our decisions take into account the representations made on all our draft determinations, responses from companies to our queries and additional information provided following further engagement with companies and other stakeholders as part of the final determination process. In the interest of brevity, where no presentations have been made on our draft determination proposals, we do not repeat our reasoning in all cases. Please see the PR19 draft determinations for further details.

For all other documents related to our final determinations, please see the final determinations webpage.
## Contents

1. Introduction .................................................................................................................................... 4  
2. Description of C-MeX and D-MeX .............................................................................................. 8  
3. Design decisions for C-MeX and D-MeX ...................................................................................... 14  
4. Design decisions for C-MeX ........................................................................................................... 26  
5. Design decisions for D-MeX ........................................................................................................... 47  
6. Next steps ...................................................................................................................................... 68  
   Annex 1: Qualitative metrics in D-MeX ......................................................................................... 69  
   Annex 2: Quantitative metrics in D-MeX ..................................................................................... 70
1. Introduction

As part of the 2019 price review (PR19) we are introducing the customer measure of experience (C-MeX) and developer services measure of experience (D-MeX). C-MeX and D-MeX are common performance commitments that are both financial and reputational. They are designed to incentivise excellent levels of service for customers in the water sector.

Both of these performance commitments are aligned to the themes of PR19 and Ofwat’s statutory duties. One of our statutory duties is to further the consumer objective to protect the interests of consumers, wherever appropriate by promoting effective competition. Customer service is one of the four key themes of PR19.

C-MeX and D-MeX take effect on 1 April 2020, when they replace the service incentive mechanism (SIM), a financial and reputational incentive which has been in place since 2010. SIM has been an effective tool and has helped water companies drive customer service improvements. However, we aim to address a number of limitations with SIM through the design of C-MeX and D-MeX:

- it was limited to comparisons within the water sector and so did not incentivise companies to reach the higher levels of customer service achieved in other sectors;
- it became apparent that there was convergence in SIM scores with diminishing improvements at the upper end, which suggested that it was not effective at encouraging leading companies to improve the customer service frontier;
- it discouraged companies from contacting their customers, which might constrain innovation, service improvement and customer engagement by water companies;
- it did not reflect changing communications technology and how customers interact with retailers in the wider economy; and
- it did not address the customer service experience of developer services customers.

1.1 Design process and timeline

In November 2016 we consulted on the future of SIM, and during 2017 we consulted on the design of C-MeX and D-MeX and made decisions on the high-level features in our PR19 methodology. Since then we have been working with customer representatives and water companies through working groups to develop and test these performance commitments.
During 2018-19, we piloted C-MeX and D-MeX (the pilot year), testing some of the components of design, and working with stakeholders to develop the mechanism and further inform the design decisions, with input and recommendations from pilot year findings.

Allto Market Research (the pilot year survey agent) was appointed as our agent to conduct the pilot year surveys and to make recommendations on survey design issues for C-MeX and D-MeX. The pilot year survey agent attended working group sessions throughout the pilot year and considered feedback from C-MeX and D-MeX working groups in finalising its recommendations.

In March 2019, we published our policy and guidance documents (the March 2019 documents) for running these performance commitments in shadow form from 1 May 2019 to 31 March 2020 (the shadow year). In some cases, we decided how C-MeX and D-MeX would operate in the shadow year, for example to test features before making final decisions for the 2020-25 period.

BMG Research (the shadow year survey agent) was appointed as our agent to conduct the shadow year surveys and to make recommendations on survey design issues for C-MeX and D-MeX. The shadow year survey agent has attended working groups throughout the shadow year and considered feedback from C-MeX and D-MeX working groups in finalising its recommendations.

Where decisions did not depend on shadow year results, we made further design decisions as part of our July 2019 draft determinations (see annex 5 of ‘PR19 draft determinations: Delivering outcomes for customers policy appendix’ from page 189).

Six companies (Bristol Water, Dŵr Cymru, Northumbrian Water, South West Water, Thames Water and United Utilities) and five other stakeholders (Business Stream, Consumer Council for Water, Fair Water Connections, Independent Water Networks and the UK Water Retailer Council) make representations in response to our draft determinations on C-MeX and D-MeX. Two of the other stakeholders (Business Stream and the UK Water Retailer Council) support C-MeX and propose similar performance commitments for business customers and retailers. We discuss these proposals in the ‘Delivering outcomes for customers policy appendix’ and the ‘Business retail market representations’.

1 See Policy decisions for the C-MeX shadow year, Guidance for the C-MeX shadow year, Policy decisions for the D-MeX shadow year and Guidance for the D-MeX shadow year.
We set out our final decisions for C-MeX and D-MeX in this document. Our final decisions take into account feedback and information from various sources, including:

- submissions made by companies and other stakeholders through representations on the draft determinations;
- feedback shared through the working group meetings over the last two years;
- recommendations made by the pilot year survey agent;
- recommendations made by the shadow year survey agent (‘BMG – Survey approach for C-MeX and D-MeX: Recommendations for the PR19 final determinations’) available on the final determinations webpage;
- recommendations from the consultants engaged by us to undertake a review of the Water UK Developer Services performance data metrics (‘Jacobs – Review of Water UK Service Levels for PR19 D-MeX Incentive’, referred to throughout this document as the Jacobs report) available on the final determinations webpage; and
- data and results from the shadow year to date.

We will publish final guidance for both performance commitments by March 2020. The next steps are further detailed in section 6.

1.2 Design principles

C-MeX and D-MeX are performance commitments to incentivise water companies to provide an excellent customer experience for their residential and developer services customers. C-MeX and D-MeX are designed to be relative performance commitments whereby a company’s performance is compared to other companies rather than against an absolute performance level. This encourages companies to innovate and improve their performance for the benefit of customers. This follows the approach used in SIM since 2010. Both performance commitments are largely survey-based, meaning that companies’ performance is principally assessed based on direct customer feedback.

C-MeX and D-MeX are both designed to have in-period financial outcome delivery incentives (ODIs) which means we will determine performance payments annually for each company during the 2020-25 period.

In designing C-MeX and D-MeX we use the following success criteria which we set out in our PR19 methodology and added to in our March 2019 documents:

- encourage companies to improve customer experiences and innovate;
be simple and meaningful for companies and customers;
be proportionate;
be practical to implement;
measure performance across companies consistently, reliably and fairly; and
reflect customer behaviour changes and market changes.

In making our decisions on C-MeX and D-MeX we balance the need for these mechanisms to be sufficiently accurate and robust to incentivise improvements in customer service, whilst minimising their complexity and cost. We seek to balance these priorities in a proportionate and practical manner.

### 1.3 Document structure

In section 2, we provide an overview of C-MeX and D-MeX, with a high-level description of the key components.

Sections 3 to 5 provide further details on key decisions made following the draft determinations.

We consider decisions that apply to both C-MeX and D-MeX (in section 3) and separately consider key decisions for C-MeX (in section 4) and D-MeX (in section 5).

We outline the next steps in section 6.
2. **Description of C-MeX and D-MeX**

This section provides an overview and high-level description of our final decisions on the key components of C-MeX and D-MeX.

Sections 3 to 5 provide further details on how we have reached our final decisions following responses to our draft determinations.

2.1 **Overview of C-MeX**

Figure 2.1 below provides an overview of C-MeX, with all its key components. Each company will have a C-MeX score calculated from responses to two surveys:

- the **customer service survey** – a customer satisfaction survey of a sample of residential customers who have contacted their company which asks them how satisfied they are with how the company has handled their issue; and
- the **customer experience survey** – a customer satisfaction survey of a randomly selected sample of a company’s overall residential customer base which asks them how satisfied they are with their company.

**Figure 2.1: Overview of C-MeX**

<table>
<thead>
<tr>
<th>Customer service survey</th>
<th>Customer experience survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>50%</td>
</tr>
</tbody>
</table>

3 points out of 100 deducted if fewer than 5 communication channels (of which 3 online channels) are offered

**Higher performance payments** apply to companies that meet all three requirements:

- One of top three performers
- Pass the cross sector upper-quartile UKCSI
- Lower than industry average on complaints

Both surveys contribute 50% each to the calculation of the overall C-MeX score for each company.

Each survey also includes a ‘likelihood to recommend’ question, the responses to which will make up the net promoter score for each company. The net promoter score will not contribute to the overall C-MeX score but will be recorded separately (see section 4.2.1).
Customer service survey approach

The customer service survey comprises monthly telephone and online interviews with customers, with annual quotas for each company of 600 telephone surveys and 200 online surveys for a total of 800 surveys per year (200 interviews per company per quarter).

For the customer service survey results, we weight the survey results for different types of contacts so that billing contributes 50% and operations 50% each. For water and wastewater companies, operations is weighted equally between water and wastewater.

Within each of these components of the customer service survey (billing, water operations and/or wastewater operations) we also weight the results of the interviews further by contact channel (by online or non-online channels) based on the relative proportion in the sample provided by companies.

We apply an upward correction factor to increase the ratings of those individual responses from customers who contact their company via online channels (see detail and reasons in section 4.2.3).

Customer experience survey approach

The customer experience survey comprises monthly telephone and face-to-face interviews with customers, with annual quotas for each company of 800 interviews (200 interviews per company per quarter).

To ensure appropriate representation of customers, we apply a quota based on national census data based on gender and age profile.

Calculating the overall C-MeX score

After applying weightings as described above, we calculate a weighted average score which represents the company’s average customer satisfaction score for each survey. Because customers provide satisfaction ratings out of 10, we rescale the scores for each survey to be out of 100 by multiplying the 0 – 10 scores by a factor of 10.

As a result each company has a customer satisfaction score for the customer service survey and customer experience survey, both out of 100. Both survey scores contribute equally to the overall C-MeX score for each company.
We then reduce a company’s overall C-MeX score by three points if it does not offer at least five communication channels, including three online channels, to receive contacts from customers (see section 4.2.2).

Each year we will publish a league table of the overall annual C-MeX scores, and a separate net promoter score, for all companies.

**Applying financial incentives**

Each company can receive outperformance payments or incur underperformance payments based on its annual C-MeX score compared to other companies. The net promoter score is not included in the C-MeX score for these purposes.

Outperformance and underperformance payments are based on where the company places relative to both the median company and either the highest or lowest performing company scores. For more detail on this, see section 3.1.

Those companies that score above the median company score will receive standard outperformance payments for that year of up to 6% of that year’s annual allowed residential retail revenue and those that score below the median company score will incur underperformance payments of up to 12%. The median company will not receive outperformance payments or incur underperformance payments.

This is demonstrated in figure 2.2 below.

**Figure 2.2: Calculations for applying financial incentives for C-MeX**

\[
\text{if } score > \text{median} : \ (score - \text{median}) \times \left(\frac{6%}{(maximum - median)}\right) \\
\text{if } score < \text{median} : \ (score - \text{median}) \times \left(\frac{12%}{(median - minimum)}\right) \\
\text{if } score = \text{median} : \ 0% \\
\]

High performing companies can also earn additional higher performance payments that increase their total outperformance payments to up to 12% of annual allowed residential retail revenue if they pass all three higher incentive gates or requirements:

- the company is one of the top three performers by C-MeX score;
- the company is at or above a cross-sector threshold of customer satisfaction performance based on the all-sector upper quartile (ASUQ) of the UK Customer Satisfaction Index (UKCSI); and
• the company has lower than the industry average number of household complaints (per 10,000 connections).

The company with the highest score that passes the three gates will receive an additional 6% of that year’s annual allowed residential retail revenue, potentially taking its total outperformance payments to 12%. If a second company qualifies, it will receive an additional 4% and if a third company qualifies it will receive an additional 2%. For the avoidance of doubt, if only one company passes the three gates it will receive an additional 6% regardless of whether it is has the highest C-MeX score across all companies.

Sections 3 and 4 provide further details on our decisions on C-MeX.

2.2 Overview of D-MeX

D-MeX is designed to incentivise water companies to provide an excellent customer experience to developer services customers, including small and large property developers, self-lay providers and those with new appointments and variations (which we refer to as ‘new appointees’). These customers can also include residential customers that have new mains connections installed.

As demonstrated in figure 2.3, D-MeX will be a single score for each company calculated based on a combination of:

• a qualitative component – a score summarising the performance of the company in a satisfaction survey of developer services customers; and
• a quantitative component – a score summarising the performance of the company across selected Water UK Developer Services performance data metrics (in this document we refer to the Water UK Developer Services performance data metrics as ‘the metrics’).

Figure 2.3: Components of D-MeX calculation
**Customer satisfaction survey approach**

For the satisfaction survey, a sample is taken of all developer services customers that transact with a company in a month. The data source for this is discussed in section 5.1.4. The sample will be based on 20% of those customers that transact with the company in a month. For each company, where 20% of the sample is expected to be more than 1,000 in a year, the sample will be capped at 1,000; where 20% of the sample is expected to be less than 100 in a year, the sample will be increased to at least 100 in a year.

Each customer surveyed gives a rating between 0 and 10 based on how satisfied they were in relation to that transaction. To calculate the qualitative component of D-MeX for each company, a simple average is taken of all responses during the year.

Some developer services customers transact with the same or multiple water companies during a year. Developer services customers will not be re-interviewed about the same company until six months after the previous interview. Where a developer services customer transacts with multiple companies they will not be interviewed about a different company until two months have passed (see section 5.1.1).

**Quantitative metrics approach**

As noted above, half of the overall D-MeX score is based on the performance of the company across selected Water UK metrics. We set out those metrics that will apply from 1 April 2020 in annex 1 of this document.

We may incorporate revised or new metrics that are developed by Water UK during the 2020-25 period. In determining whether to make changes to D-MeX as a performance commitment in light of changes to the Water UK metrics, our decision will be based on the principles that revisions are in customer interests, support consistent and fair comparisons between companies and align with our wider duties. We will consult with stakeholders prior to making any changes.

To calculate a company’s quantitative component of D-MeX, a simple average is taken of the scores applying to each metric for that year. Metrics which do not have activity recorded against them are excluded from the calculation for that company.

**Calculating the D-MeX score**
As we say above, for each company a simple average is taken of customer satisfaction ratings. Because customers provide satisfaction ratings out of 10, we rescale this average score to be out of 100 by multiplying it by a factor of 10.

A company’s score for the quantitative component will be reported out of 100.

Each component contributes 50% of the overall D-MeX score. Each year we will publish a league table of the overall annual D-MeX scores for all companies.

**Applying financial incentives**

Each company can receive outperformance payments or incur underperformance payments based on its annual D-MeX score compared to other companies.

Outperformance and underperformance payments are based on where the company places relative to both the median company and either the highest or lowest performing company scores. For more detail on this, see section 3.1.

Those companies that score above the median company score will receive standard outperformance payments for that year of up to 6% of that year’s annual allowed residential retail revenue and those that score below the median company score will incur underperformance payments of up to 12%. The median company will not receive outperformance payments or incur underperformance payments.

This is demonstrated in figure 2.4 below.

**Figure 2.4: Calculations for applying financial incentives for D-MeX**

\[
\begin{align*}
\text{if score} & > \text{median} : (\text{score} - \text{median}) \times (6\%)/(\text{maximum} - \text{median}) \\
\text{if score} & < \text{median} : (\text{score} - \text{median}) \times (12\%)/(\text{median} - \text{minimum}) \\
\text{if score} & = \text{median} : 0\% 
\end{align*}
\]

Sections 3 and 5 provide further details on our decisions for D-MeX.
3. Design decisions for C-MeX and D-MeX

In this section we discuss design decisions that we consider apply to both C-MeX and D-MeX.

In the interests of simplicity and ease of understanding for customers and companies, we align design features of C-MeX and D-MeX where it is appropriate to do so.

3.1 Application of financial incentives

3.1.1 What we said in our draft determinations

In our draft determinations we set out approaches in C-MeX and D-MeX to allocate financial incentives to companies. For C-MeX, we proposed that standard outperformance and underperformance payments for each year would vary between +6% and -12% of annual allowed residential retail revenue for that year, based on how far a company’s score is from the average score (i.e. the level of standard deviations from the mean) with the following ‘steps’ or payment bands:

- for scores between 0 and ±0.25 standard deviations from the mean average a deadband or 0% payments will apply;
- for scores between +0.25 and +1 standard deviations above the mean average, the outperformance payments will be the company’s standard deviations from the mean multiplied by 3%;
- for scores between -0.25 and -1 standard deviations below the mean average, the underperformance payments will be the company’s standard deviations from the mean multiplied by 6%; and
- for scores ±1 standard deviations from the mean the highest outperformance payments (+6%) or underperformance payments (-12%) will apply.

For D-MeX we proposed this would be based on the quartile in which companies place in the D-MeX league table, so that:

- for companies above the upper quartile, outperformance payments of 6% will apply;
- for companies between the upper quartile and median, no payments will apply;
- for companies between the median and lower quartile, underperformance payments of -6% will apply; and
• for companies below the lower quartile, underperformance payments of -12% will apply.

We considered that different approaches for C-MeX and D-MeX were appropriate – for C-MeX we largely followed the approach set out in our PR14 final determinations for SIM relating to performance in the 2010-15 period; for D-MeX we adopted the quartiles approach because we observed greater outliers in the pilot year for D-MeX than C-MeX, and the quartiles approach limits the impact of extreme outliers on other companies’ scores.

3.1.2 Stakeholders’ representations

United Utilities states that the current design for C-MeX is likely to drive inappropriate outperformance and underperformance payments which will undermine the integrity of C-MeX. The company proposes we remove ‘cliff edges’ from the incentive calculations where companies can receive highly different payments for minor differences in scores; it claims a more gradual approach to calculating incentive payments would be preferable - to ensure companies with similar scores receive similar payments.

Northumbrian Water states that using standard deviations from the mean to apply financial incentives for C-MeX could result in a single significant underperforming company reducing outperformance payments for other companies. The company also states that, given the complexity of the approach, further consultation would be beneficial.

Bristol Water states the proposed quartiles approach for D-MeX is disproportionate because it can lead to high variation in payments between quartiles. The company suggests we use the approach we proposed to use for C-MeX.

Dŵr Cymru states the proposed approach for D-MeX is not reasonable because it could lead to mid-performing companies receiving underperformance payments. The company claims no penalties should apply for the middle two quartiles, at least until it becomes clear whether the D-MeX scores will be stable or volatile for companies.
3.1.3 Our assessment and reasons

C-MeX considerations

Based on shadow year data compared to the pilot year, we observe a higher risk of significantly different payments for companies with statistically similar scores, for example where companies are close to the ‘steps’ that are inherent in the approach. For example, a step can be seen to exist where payments increase from nearly +3% if a company is close to +1 standard deviation from the mean, and increase to +6% at or above +1 standard deviation from the mean. While this approach introduced additional incentives to avoid very high or very low performance because of these cliff edges, we recognise that the way we initially proposed to apply incentives could lead to companies with similar scores receiving significantly different payments.

We decide to adopt a more linear approach that applies incentive rates based on how far a company’s score is from the highest and lowest performing companies, which receive +6% and -12% respectively. We also aim to ensure the average performer does not receive underperformance or outperformance payments and to reduce the impact of outlier companies. We decide to anchor the calculation around the median company. This means a company’s payments will instead be based on what its C-MeX score is, relative to the median company and either the highest or lowest performing company. Importantly, this decision and our decision to use the median instead of the mean further reduces the impact of negative outliers on other companies, particularly for companies above the median that are due to receive outperformance payments.

We illustrate how this approach works in figure 3.1 using C-MeX scores from the shadow year.
This approach ensures payments are significantly more reflective of a company’s performance relative to other companies. It is also less complex and therefore easier to communicate than the approach set out at draft determinations.

Because our final approach removes these steps for calculating payments, it could be argued that it weakens incentives on companies to achieve very high performance or avoid very low performance. However, significant incentives remain because there are continuous incentives to improve due to the lack of deadbands and due to other design features of C-MeX such as its relative nature, the higher payments threshold and the reputational league table.

Our final approach is in line with representations to our draft determinations (to remove cliff edges) as it more closely aligns payments with scores.

While our final approach is similar to how we reconcile SIM at PR19 (see ‘Accounting for past delivery technical appendix’) we consider it appropriate to use a modified approach for C-MeX and D-MeX to reflect that we are setting an incentive framework for company performance in the upcoming 2020-25 period, whereas how we decide to reconcile SIM at PR19 is for past performance and so is not intended to affect future company behaviour. Furthermore, in the pilot and shadow years of C-MeX and D-MeX we identify a greater extent of negative outlier companies than
under SIM at PR19 which could lead to a majority of companies receiving outperformance payments.

This approach has a minor implication for how we award higher outperformance payments for C-MeX, which we set out in section 4.4.

**D-MeX considerations**

At draft determinations, we applied the quartiles approach for D-MeX because we observed greater outliers in the pilot year for D-MeX than C-MeX, and the quartiles approach limits the impact of extreme outliers on other companies’ scores. However the shadow year results do not show the same spread of scores or outliers for D-MeX as were observed in the pilot year.

We recognise that significant steps in payment amounts can occur between quartiles despite similar scores. For example, based on shadow year data it is possible that under the quartiles approach one company receives +6% outperformance payments but scores only 0.9 points above another company that faces -6% underperformance penalties.

We do not agree with Dŵr Cymru’s proposal to apply a deadband for the middle two quartiles of companies, as this provides insufficient incentive for companies to improve.

In the interests of simplicity and ease of understanding for customers and companies, we align features of C-MeX and D-MeX where appropriate.

We observe a similar spread of scores and outliers between C-MeX and D-MeX scores in the shadow year, strengthening the case to apply a common approach to both performance commitments.

**3.1.4 Our final determination decision**

Following our assessment, we change our decision for the final determinations. For both C-MeX and D-MeX we apply financial incentives for companies based on where they place relative to the median company and either the highest or lowest performing company. This is demonstrated in figure 3.2 below.
Figure 3.2: Calculations for applying financial incentives for C-MeX and D-MeX

\[
\begin{align*}
\text{if score > median : } & (score - \text{median}) \times (6\% / (\text{maximum} - \text{median})) \\
\text{if score < median : } & (score - \text{median}) \times (12\% / (\text{median} - \text{minimum})) \\
\text{if score = median : } & 0\%
\end{align*}
\]

As part of our annual reconciliation of in-period outcome delivery incentives, each year we apply the incentive rate calculated in this way to the company’s annual allowed residential retail revenue for C-MeX for that year, and the company’s annual actual developer services revenue for that year (as defined in section 5.3.4) for D-MeX.

3.2 Overall design issues

In this section we discuss four design issues raised by stakeholders:

- absolute rather than relative measures;
- customer willingness to pay;
- grouping companies by performance bands; and
- company-specific targets.

3.2.1 What we said in our draft determinations

In our draft determinations we said C-MeX and D-MeX would be relative incentive mechanisms that measure the performance of companies relative to each other. We made no comment on company specific targets or customer willingness to pay.

3.2.2 Stakeholders’ representations

Thames Water proposes an approach to C-MeX where companies receive payments related to absolute improvements rather than their relative performance to other companies. The company states relative scoring does not recognise improvements in performance, or the relative differences between companies and their customers’ expectations. The company claims that, under our approach proposed in our draft determinations, underperformance payments will be paid by the worst performing companies even if they significantly improve their performance.
Thames Water also states C-MeX does not take into account what customers are willing to pay for customer service, and as such the incentives are not linked to benefits generated for customers.

United Utilities proposes grouping companies into performance bands (such as good performers, average performers or poor performers) would be more appropriate given the uncertainty of data (rather than ranking individual companies 1 to 17 in end-of-year league tables).

The Consumer Council for Water proposes we set company-specific targets for C-MeX scores to drive further improvements. In particular, the Consumer Council for Water proposes Bristol Water receives a specific complaints performance target as part of its C-MeX performance commitment to help drive further improvements.

### 3.2.3 Our assessment and reasons

In relation to whether C-MeX should be an absolute measure, while there are benefits in using either an absolute or relative measure, we set out in our PR19 methodology that C-MeX will be a relative measure of performance, which has been the approach to SIM since 2010. SIM has proven to be effective in measuring companies’ performance relative to other companies and encouraging them to continually innovate and improve performance. In particular, we consider companies should receive outperformance payments for customer service performance that is comparatively better relative to other companies, and a relative measure delivers this effectively. The relative nature also provides strong incentives for companies to improve relative to other companies.

**Customers’ willingness to pay** studies are used to estimate how much customers value improvements in performance levels and are an input when we set outcome delivery incentive rates. However C-MeX and D-MeX are different to other outcome delivery incentives because they are relative measures that focus on driving improvements relative to all other companies, rather than incentivising companies to achieve specific service levels. Because C-MeX and D-MeX are relative measures it is not appropriate to link outperformance or underperformance payments to willingness to pay in the same way as other outcome delivery incentives.

Regarding **grouping instead of ranking companies**, we consider doing so would substantially weaken reputational incentives for companies to improve their performance relative to other companies and could confuse customers and other stakeholders when comparing each company’s C-MeX score. Publishing league tables for all companies maintains a strong incentive for companies to improve their
position throughout the 2020-25 period, as we have observed in the 2015-20 period under SIM. We note that this representation from United Utilities is at least in part based on concerns about accuracy of scores, which we address in section 4.1.

In terms of company-specific targets, we do not set company-specific targets within C-MeX because it is a relative measure and doing so may add unnecessary complexity. Companies are also able to set their own targets relating to customer service. For these reasons we do not consider it appropriate to set further company-specific targets within C-MeX or D-MeX.

Taking into account all of these issues, we consider that C-MeX and D-MeX are most appropriately structured as performance commitments that encourage companies to improve their performance relative to other companies.

### 3.2.4 Our final determination decision

We maintain our decisions from our draft determinations. C-MeX and D-MeX are relative common performance commitments based on rankings, with no performance bands or company-specific targets.

### 3.3 Weighting of water and wastewater scores

#### 3.3.1 What we said in our draft determinations

In our draft determinations, we acknowledged concerns that water only companies could be at a disadvantage compared to water and wastewater companies. Based on advice from the pilot year survey agent, we found too small a difference between water and wastewater company scores in C-MeX and D-MeX to affect rankings or financial incentives.

In the March 2019 documents and our draft determinations, we stated we would apply an equal weight to water and wastewater in both C-MeX and D-MeX.

The pilot year survey agent considered options for weighting water and wastewater and noted the concern expressed by water only companies that equal weighting was unfair. The concerns of water only companies were twofold: first, wastewater satisfaction scores are higher than water scores, and secondly, equal weighting of water and wastewater has the effect of attributing a higher weight to wastewater than water, as there are around twice as many water related contacts from customers.
The pilot year survey agent noted that adopting weights based on the proportion of contacts reduced the customer service survey satisfaction scores for water and wastewater companies slightly, with minimal impact on ranking.

### 3.3.2 Stakeholders’ representations

Bristol Water states that it is unable to confirm if it is comfortable with the weightings applied to water and wastewater before it has seen a breakdown of the shadow year C-MeX results. Bristol Water notes that wastewater satisfaction scores have further improved since private sewer adoption and suggests weightings should be adjusted so the impact of water and wastewater satisfaction scores is on average the same within C-MeX. Bristol Water proposes a new correction factor in line with the online correction factor already used in C-MeX.

Bristol Water agrees with an approach that includes both water only companies and water and wastewater companies in the same league table for D-MeX, subject to using a gradual approach to applying financial incentives. Otherwise, it considers our approach potentially disproportionate to the level of risk.

### 3.3.3 Our assessment and reasons

We analyse any differences that exist in scores between water only and water and wastewater companies, and between water and wastewater satisfaction scores across companies. For C-MeX, we use the last ten years of SIM scores and C-MeX data from the shadow and pilot years. For D-MeX, we analyse D-MeX pilot year data.

For the scores **between water only companies and water and wastewater companies** we observe from SIM data that water only companies have historically performed better than water and wastewater companies, but since 2016 water and wastewater companies have improved their performance and outperformed water only companies. However, when considering the trends over time, and the data observed in the C-MeX pilot and shadow years, we see no conclusive statistical evidence of any difference in scores between water only and water and wastewater companies.

This evidence therefore provides no justification for different weightings, a correction factor, or a separate league table for water only companies separate from water and wastewater companies.
We also analyse the underlying water and wastewater scores **within water and wastewater companies** to see if there are any statistically significant differences between the two operational business units. Ten years of SIM data show that, on average, wastewater scores have been approximately 2% higher than water scores; since 2017 this has increased to over 4%\(^2\). While this difference is statistically significant for SIM, analysis of data from the shadow year is unable to identify any statistically significant difference between water and wastewater scores within water and wastewater companies. Because the evidence is not consistent, there is no justification for changing the weighting of water and wastewater scores within a water and wastewater company.

Further, we are concerned that reducing the weight placed on wastewater services within a water and wastewater company would weaken the incentive for them to continue to improve wastewater customer service, failures of which can have a significant impact on customers. For these reasons, we maintain our approach of equal weightings between water and wastewater contacts (for water and wastewater companies) within the operational component of the customer service survey.

Compared to SIM, C-MeX already places less weight on wastewater services, as satisfaction scores in the customer service survey contribute 50% of the score compared to 75% in SIM. In addition, the C-MeX customer experience survey focuses on customers’ experience with their water only company where they receive water and wastewater services from different companies. The implication of this is that there may be fewer wastewater responses (than water responses) in the calculation.

Based on pilot year results for D-MeX, we find evidence that wastewater services perform slightly better than water services in both the qualitative and quantitative component of D-MeX. However, as with C-MeX, the difference is too small to impact D-MeX rankings.

### 3.3.4 Our final determination decision

Following our assessment, we maintain the approach set out in our draft determinations for both C-MeX and D-MeX. We continue to include both water only and water and wastewater companies in a single league table, and the water and wastewater activities within water and wastewater companies continue to receive the same weighting in the operational component of the customer service survey. We

\(^2\) This was not true for two companies: for one company, the difference between water and wastewater services decreased; for another company, its water services scored higher than its wastewater services throughout, although this difference has reduced in the last two years.
will monitor to see if the trend continues over the 2020-25 period to inform decisions in the next price review.

Bristol Water’s concerns in relation to D-MeX are mainly addressed in our analysis of the application of financial incentives discussed in detail in section 3.1.

### 3.4 Data assurance

#### 3.4.1 What we said in draft determinations

In the March 2019 documents, we noted that when D-MeX becomes operational we expect companies to provide us with data for the quantitative element of D-MeX that has gone through sufficient assurance processes, similar to the data provided for other performance commitments.

In our draft determinations we welcomed companies’ feedback on the guidance and said that we are interested in how companies propose to provide assurance.

#### 3.4.2 Stakeholders’ representations

Fair Water Connections considers that there are many anomalies in company reporting in relation to the Water UK Levels of Service metrics.

#### 3.4.3 Our assessment and reasons

We acknowledge that the level of assurance on data underlying the selected set of Water UK metrics in D-MeX needs to be sufficient and proportionate, particularly because it will be a new measure of performance from April 2020.

We expect companies to report actual performance against these measures in their annual performance reports from 2020-21. We expect companies to provide commentary to outline their assurance process and the findings. In line with our policy for other performance commitments, we note that it is each company’s responsibility to report accurate and complete information and it must have adequate processes in place to do this.

For C-MeX, companies are required to offer at least five communication channels for receiving contacts and complaints, three of which must be online, throughout the
reporting year. A reduction of three points out of 100 will be applied if the requisite number of channels is not offered. To enable this adjustment to be applied, companies should provide an assurance statement in their annual performance report that the requisite number of channels have been offered throughout the reporting year.

3.4.4 Our final determination decision

Companies are required to provide assurance in their annual performance reports for all performance commitments. For C-MeX and D-MeX, we require the following additional reporting and assurance requirements which are reflected in the ‘Outcomes performance commitment appendix’ for each company:

- their performance against the metrics as reported to Water UK for the relevant year;
- a statement that confirms whether the company offered at least five communication channels for receiving customer contacts and complaints and at least three online channels throughout the reporting year; and
- the process the company has taken to assure itself that its performance against the selected Water UK metrics in D-MeX are an accurate reflection of its underlying performance in the reporting year, as well as any findings that indicate this is not the case.
4. Design decisions for C-MeX

In this section we consider stakeholder representations and design issues that are specific to C-MeX.

4.1 Accuracy of C-MeX scores

There are several factors that contribute to the accuracy of C-MeX scores: sample size; the relationship between the net promoter score question and the customer satisfaction question; how quotas and weightings are applied; and the frequency with which survey results are calculated. These factors are all interrelated, and the way they work together in C-MeX also affects the accuracy of scores.

4.1.1 What we said in our draft determinations

In our PR19 methodology we said C-MeX would be based on an annual survey of 200 interviews per survey per company per quarter – or 800 interviews per survey per company per year or 1,600 interviews per company per year (both surveys combined). We said we would look at sample sizes in the pilot year. In our draft determinations we said this sample size was appropriate because it provides a robust result at a reasonable cost and is in line with the recommendation of the pilot year survey agent.

In the March 2019 documents we also confirmed our intention to include the net promoter score (NPS) in the C-MeX calculation alongside the customer satisfaction scores from the customer service survey (CSS) and customer experience survey (CES):

\[
C\text{-}MeX \text{ score } = 40\% \text{ CSS-CSAT} + 40\% \text{ CES-CSAT} + 20\% \text{ NPS (from CSS and CES)}
\]

We said the net promoter score is a strong comparative tool but noted the pilot year data showed poor customer understanding of the net promoter score question relative to the customer satisfaction question, and that there was volatility in the pilot year. We said we would monitor the volatility and understanding of the net promoter score in the shadow year.
4.1.2 Stakeholders’ representations

United Utilities states the current design and sample sizes proposed for C-MeX surveys are likely to result in a high degree of incorrect relative company rankings year on year, directly driving inappropriate outperformance and underperformance payments. The company states such errors undermine the integrity of a high-profile performance commitment.

United Utilities’ representation is supported by a report from Frontier Economics that assesses the statistical validity of C-MeX. In summary Frontier Economics finds:

- given the sample size for C-MeX, there may not be a real performance difference between companies with annual scores less than 2.7 points apart;
- companies may not be ranked correctly relative to each other – under the current design companies have a 40% to 75% chance of being ranked incorrectly;
- under the manner in which incentive payments are proposed to be applied, companies have a 11% to 37% chance of being in the wrong payment ‘bucket’ – and this will not on average cancel out due to the asymmetric design of financial incentives and the presence of ‘cliff edges’;
- the presence of the net promoter score in the C-MeX calculation adds to the statistical ‘noise’ in the overall score and is highly correlated with other C-MeX components; and
- reallocating sample size from the customer experience survey to the customer service survey and other components of C-MeX can create a more precise score.

Frontier Economics bases its analysis on data for United Utilities and suggests we replicate some of the statistical tests it has used with data for all companies and the shadow year.

United Utilities proposes a number of steps to overcome these issues and improve the statistical accuracy of and stakeholder confidence in C-MeX:

- at least double C-MeX sample sizes to increase confidence intervals;
- remove the net promoter score from the C-MeX calculation to boost statistical significance of company scores;
- consider how the sample size is allocated between the customer service and customer experience surveys of C-MeX; and
- remove ‘cliff edges’ from the way incentive payments are calculated.

No other stakeholders comment specifically on statistical confidence in C-MeX but some stakeholders comment on how we apply financial incentives or the design and calculation of C-MeX. Northumbrian Water comments on the impact of outliers on
incentive payments while Thames Water proposes simplifying C-MeX by removing either the net promoter score or the customer satisfaction score from the C-MeX calculation. We address both points in sections 3.1 and 4.2.

4.1.3 Our assessment and reasons

We recognise that wide confidence intervals around a company's score may lead to rankings or incentive payments that do not accurately reflect a company's relative performance. We are mindful that our design for calculating C-MeX scores should produce results sufficiently robust to incentivise improvements in customer service, while remaining proportionate, cost effective and not too complex for stakeholders to engage with.

We recognise some of our policy design decisions impact the level of statistical accuracy in companies' scores. We can also mitigate the impact by revising how we apply financial incentives.

We assess the representation from United Utilities and repeat some of the statistical tests using pilot and shadow year data for all companies. As noted above, there are four main factors impacting on accuracy. We address these in turn:

- our approach to quotas and weightings (in this section);
- our approach to calculating annual scores (in this section);
- an appropriate sample size for C-MeX surveys (in this section); and
- use of the net promoter score in C-MeX (in section 4.2.1).

Further, the way in which we apply financial incentives can mitigate the impact on companies of inaccurate scores (see section 3.1).

Our approach to quotas and weightings

We set quotas (or targets) for customer types to ensure appropriate representation across customers and areas of service delivery in the survey sample. For example in the customer service survey, we specify a percentage of interviews must be from customers that contacted their company about operational as opposed to billing issues. This helps to achieve a representative sample of customer experience.

Once the interviews are complete, to arrive at the overall C-MeX score we also apply weightings to responses based on customer type and whether contact was made by an online channel. Weightings are important to ensure that there is adequate representation of those customer types – for example in the customer service survey
billing contacts would generally dominate operational contacts, and within operations water contacts would dominate wastewater contacts. Without weightings there is a risk companies do not focus sufficiently on these customers or business areas.

For the shadow year, the way we applied quotas and weightings were not aligned. For instance we applied quotas of 33% billing, 33% water operations and 33% wastewater operations for water and wastewater companies. A further weighting is applied within these customer type weightings to reflect the proportion that contacted their company via online channels. When this 33-33-33 quota is combined with a 50-25-25 weighting, the difference between them has the effect of reducing the effective sample size and statistical confidence of the survey results.

The shadow year survey agent recommends aligning quotas to be equal to the weightings for each survey. This reduces the level of difference between them and therefore has the effect of increasing the effective sample size.

We consider this recommendation by the shadow year survey agent to be a reasonable and cost-effective way to help improve the overall accuracy of C-MeX scores. By aligning quotas to weightings we increase the effective sample size and improve overall confidence in C-MeX. In this way, we apply quotas aligned with the weightings we set out in section 2.1.

**Our approach to calculating annual scores**

In our draft determinations, we said we would apply C-MeX in-period on the basis of an average of four quarterly scores.

Calculating annual C-MeX scores based on annual data instead of an average of four quarters means the results are based on a larger sample size – 800 per company per survey, as opposed to 200 per company per survey in a quarter. All other things being equal, this gives a more accurate score. We consider the improvement in the accuracy of C-MeX scores to be desirable for the purposes of applying in-period financial incentives.

We recognise that companies are interested to track their performance throughout the year. We will therefore provide interim scores to companies every quarter, with weightings applied on a rolling basis, which means that the first quarter will be based on data and weightings for three months of the year, the second quarter on six months of the year and so on. This allows companies to track their progress towards their annual score, and encourages them to improve their performance during the course of the year.
While we will provide quarterly scores to companies, for the reasons set out above, we consider it appropriate to calculate overall C-MeX scores on an annual basis rather than as an average of four quarterly scores.

**An appropriate sample size for C-MeX**

While increased sample size can lead to improved accuracy of scores, there are diminishing returns from increasing sample size. Further, this has to be considered in light of the survey costs – increasing sample size is costly, particularly in the customer experience survey, where face-to-face interviews are increasingly relied on due to the low response rates and declining coverage of telephone interviews. Given the other changes we make to improve accuracy of C-MeX scores, we consider that the sample size used in the shadow year is appropriate for the 2020-25 period.

**How we apply financial incentives**

We agree that how we apply financial incentives can exacerbate the impact of less accurate scores. We set out our revised approach in section 3.1, but in summary we decide to apply a more gradual approach to calculating incentive payments to ensure companies with similar scores receive similar payments. Based on our assessment this substantially mitigates the financial impact of statistical confidence issues.

**4.1.4 Our final determination decision**

To help improve the overall accuracy of C-MeX scores, in summary we decide to:

- align all quotas with weightings for each subcomponent of C-MeX (discussed above in this section);
- change how we calculate company scores to be based on annual results instead of an average of four quarters (discussed above in this section); and
- remove the net promoter score from the C-MeX calculation (which we discuss in the following section, 4.2.1).

We retain the current sample size (discussed above in this section). We change the way we apply financial incentives (discussed in section 3.1 above).
4.2 How we calculate C-MeX scores

4.2.1 Net promoter score

What we said in our draft determinations

The net promoter score is a customer experience metric developed as an alternative to customer satisfaction. It is calculated from responses to a question (typically ‘Based on your recent experience, how likely would you be to recommend this company to friends or family?’) using a scale of 0 to 10. Individual ratings are categorised into detractors (0 to 6), passives (7 or 8) or promoters (9 or 10), with the net promoter score being the proportion of promoters minus detractors.

In our draft determinations we included the net promoter score directly in the C-MeX calculation with customer satisfaction (CSAT) scores from the customer service survey and customer experience survey as follows:

\[ C\text{-MeX score} = 40\% \, CSS\text{-CSAT} + 40\% \, CES\text{-CSAT} + 20\% \, NPS \text{ (from CSS and CES)} \]

We said the net promoter score is a strong comparative tool and easy for companies to understand but noted the pilot year data showed poor survey respondent understanding of the net promoter score question, as compared with the customer satisfaction question, and that there was volatility shown in net promoter scores across the pilot year. We said we would monitor the volatility and understanding of the net promoter score in the shadow year.

Stakeholders’ representations

As discussed in section 4.1, United Utilities proposes removing the net promoter score from the C-MeX calculation to increase the accuracy of C-MeX scores and because customers do not understand how the net promoter score differs from general satisfaction ratings.

Frontier Economics, commissioned by United Utilities, estimates over half of the variance in C-MeX arises from the net promoter score. It estimates that removing the net promoter score would have the same effect on statistical confidence as almost doubling the sample size.

Thames Water proposes simplifying C-MeX by basing it on a single question, either the net promoter score or customer satisfaction question (for both surveys). Thames
Water states the combined customer satisfaction and net promoter score survey has become complex and hard to explain to customers and employees.

Bristol Water states how we calculate the net promoter score does not follow the standard methodology. It also proposes publishing a separate league table for the net promoter score to understand how it affects overall C-MeX rankings.

**Our assessment and reasons**

The net promoter score is an acknowledged measure of overall customer experience with a company’s product service. It can provide a form of cross-sector comparison and has been used to stimulate competition in some regulatory frameworks by showing relative scores between service providers. However there is some evidence that including the net promoter score as a component of the C-MeX calculation is reducing the overall accuracy of C-MeX scores.

We assess how to use the net promoter score in C-MeX:

- The net promoter score can introduce a cross-sector comparison because it may encourage respondents to think about their water company in the same way they would think about providers in other sectors.
- Using shadow year data, we observe volatility in net promoter scores, sometimes without obvious explanation. This is partly a consequence of how the net promoter score is calculated – customers are asked to provide a rating between 0 and 10, which is then converted to a three-point scale (promoters, detractors, and passive customers), before being rebased to a score out of 100. However, the shadow year survey agent observes that this volatility has remained at similar levels in the shadow year in both the customer service and customer experience surveys, and that these levels are acceptable.
- We find customer understanding is broadly unchanged from the pilot year results. The pilot year survey agent found that 16% of respondents could not respond to the net promoter score question compared to around 1% for the customer satisfaction question. The shadow year survey agent reports that around half of respondents to the customer experience survey, and one in ten respondents to the customer satisfaction survey, were confused with the question. Even after prompting or explanation, 4% of respondents to the customer service survey and 14% of respondents to the customer experience survey were unable to provide a response to the net promoter score question.
- Our analysis suggests that the net promoter score question adds little to the C-MeX score because of its relatively high correlation with responses to the customer satisfaction question in both surveys (the customer service and customer experience surveys). That is, customer responses to the net promoter
score question typically align with their responses to the customer satisfaction questions for both surveys. Because they are aligned, adding the net promoter score adds little new information while also adding statistical noise. This widens the confidence intervals around a company’s C-MeX score, making them less accurate.

- The net promoter score increases divergence between companies’ C-MeX scores. However, we observe that much of this divergence is a product of the less accurate net promoter scores. That is, the observable divergence is not necessarily a result of any additional information to distinguish companies’ performance.

As a result we consider including one but not both of the customer satisfaction and net promoter scores in each survey. C-MeX is designed as a measure of customer experience. Responses to the customer satisfaction questions in the surveys provide a more accurate measure of a customer’s experience than responses to the net promoter question. For this reason we decide to base the customer service survey and customer experience survey scores on the customer satisfaction question alone.

We recognise C-MeX is a new measure for PR19 and more data may be needed to assess the impact of the net promoter score relative to a measure of customer satisfaction. Net promoter scores will still be recorded and published in an annual league table but do not contribute to the calculation of the C-MeX score which financial incentives will be based on.

Highlighting the net promoter score in this way provides information to customers and other stakeholders to enable direct comparison of company performance under both measures. It also reduces the complexity of the C-MeX calculation, and provides five years of data comparing C-MeX and net promoter scores to inform decisions for the 2024 price review (PR24).

**Our final determination decision**

Our final decision is to remove the net promoter score from the calculation of C-MeX financial incentives. C-MeX scores for each company will be calculated as follows:

\[
C-MeX\ score = 50\% \ CSS-CSAT + 50\% \ CES-CSAT
\]

Net promoter scores will still be recorded and separately to C-MeX scores without contributing to the calculation of financial incentives.
4.2.2 Communication channels adjustment

What we said in our draft determinations

We said we would apply a three-point downward adjustment (out of 100) to a company’s overall C-MeX score if it does not offer at least five communication channels, including at least three online channels, to receive customer contacts and complaints.

Stakeholders’ representations

Bristol Water proposes that the objective behind the communication channels adjustment may be better enacted as a formal regulatory commitment rather than through C-MeX due to the complexity it adds to the C-MeX calculation.

Our assessment and reasons

We consider it important for customers to be able to contact their company through a variety of communication channels. One of the shortcomings identified with SIM was that it did not recognise changing communication channels.

Based on shadow year data, we can see companies submitted contact data from at least five different channels. Only one company lacks three online channels.

While we recognise this approach adds an element of complexity, it does provide a notable incentive on companies, reflects customer behaviour changes, and recognises the changing nature of channels required to communicate with customers, particularly online channels. We consider this is proportionate given shadow year data indicates limited improvements are required by companies.

Companies will be required to provide assurance in their annual performance reports that they have the requisite number of communication channels offered to customers.

Our final determination decision

We maintain our decision from draft determinations. We will apply a downward adjustment to a company’s overall C-MeX score if it does not offer at least five communication channels, including at least three online channels. For clarity, this will be a reduction of three points from the score (which is out of 100), not a 3% reduction in score.
4.2.3 Online correction factor

What we said in our draft determinations

We said we would apply an upward correction factor of 5% to individual survey results from customers who responded online to the customer service survey (for instance a score of 8 out of 10 would be multiplied by 105% to become 8.4). This was in recognition that they are significantly less positive than phone survey results according to pilot survey data and feedback from companies.

We asked the shadow year survey agent to test this in the shadow year.

Stakeholders’ representations

Bristol Water states it is comfortable with the approach of applying a correction factor but cannot provide more conclusive comments due to limited comparison data provided. The company asks why we reduced the factor from 6% to 5%.

Our assessment and reasons

We assess the appropriateness of the level of online correction factor, which was initially based on the recommendations made by the pilot year survey agent as a result of the sampling bias that was observed from online surveys being noticeably less positive than those of the telephone surveys.

In our draft determinations we proposed to adopt an online correction factor of 5% (not 6%), based on samples and results that the pilot year survey agent had used to make its final recommendation.

We further test the online correction factor in the shadow year, whereby the shadow year survey agent conducts a desktop research exercise to examine general differences between telephone and online surveys. In its recommendations the shadow year survey agent concludes there is no evidence to challenge the recommendations from the pilot year survey agent.

We decide to apply an online correction factor that is reasonable and consistent with the shadow year survey agent’s findings. Doing this ensures that companies with a higher proportion of digital contacts do not receive a lower C-MeX score than those companies with a lower proportion of digital contacts due to this sampling bias.
Our final determination decision

We maintain our decision from our draft determinations to apply an online correction factor of 5% to individual survey responses from online channels to the customer service survey.

4.2.4 Socio-demographic factors

What we said in our draft determinations

We said we would not impose a further quota or weighting for customers’ socio-economic (or socio-demographic) status given the increased complexity and cost this would add to the survey, and the absence of statistical difference between the results.

The pilot year survey agent did not recommend a quota because it found no statistically significant difference and said it would increase the difficulty of achieving the interviews and the cost of the survey.

Stakeholders’ representations

Thames Water proposes that different company results are not comparable due to differences between each company with regard to socio-demographics of customers, i.e. relating to the specific needs, challenges, priorities and customers of the region the company serves. Thames Water states such differences could be addressed through weighting while noting this would add some complexity.

Thames Water references a 2016 report submitted as part of their 2018 business plan submission that analyses the association between socio-demographic factors and customer service scores. This report provides a mixture of results, including both statistically significant and insignificant associations.

Our assessment and reasons

All companies are challenged by the socio-demographic features of their customer base. This requires companies to understand their customers and respond accordingly.

If we account for socio-demographic differences in C-MeX, we must take account of the customer profile of each company in any weighting or correction factor. This could change every year. Accounting for this would add a degree of undue
complexity to C-MeX. This complexity is acknowledged by Thames Water in its representation, and was also identified by the pilot year survey agent.

Furthermore, the pilot year survey agent observed no statistically significant socio-demographic differences between companies during the pilot year that warrants weightings or a correction factor.

**Our final determination decision**

We maintain our decision in our draft determinations. We do not correct for the socio-demographic features of companies' customers (other than apply a quota for age groups, see section 4.3.3) because the pilot year survey agent and the submission from Thames Water find mixed evidence of an impact of socio-economic demographics on C-MeX scores. Without clear evidence, further weightings or correction factors are not justified. Adding them would add to the complexity of C-MeX without clear benefit.

**4.2.5 Treatment of Hafren Dyfrdwy in the customer service survey**

**What we said in our draft determinations**

For the customer service survey we said we would combine the operational wastewater results for Severn Trent Water and Hafren Dyfrdwy to produce one wastewater score for both companies.

**Stakeholders’ representations**

No stakeholders comment on this issue.

**Our assessment and reasons**

Given the size of its wastewater business, Hafren Dyfrdwy is likely to have a very small sample size for wastewater results from the customer service survey. This continues to be the case based on shadow year results.

We continue to consider merging the wastewater results for Hafren Dyfrdwy and Severn Trent Water is a proportionate and practical approach to the issue, as the companies share a call centre and wastewater contractors, and the wastewater results in the customer service survey is a relatively small part of the overall C-MeX score.
Our final determination decision

We decide to continue to merge the wastewater results for Hafren Dyfrdwy and Severn Trent Water in the customer service survey given the very small sample size for Hafren Dyfrdwy’s wastewater results.

4.3 C-MeX surveys

In our draft determinations we said there were a number of survey-related issues to be monitored and reviewed before survey methodologies are finalised.

4.3.1 Sampling approach for the customer experience survey

What we said in our draft determinations

In the draft determinations we said that we adopt a predominantly telephone-based survey using a random digit dial, supplemented by face-to-face interviews to ensure sufficient interviews are achieved with younger age-groups.

Stakeholders’ representations

Bristol Water states it cannot support what it says is a significant change in approach for the pilot year. The company states the customer experience survey sample is now 50% face-to-face and 50% telephone surveys.

The company is concerned using face-to-face interviews will not take place in geographic areas that are representative of a company’s customer base.

Our assessment and reasons

Based on recommendations made by the pilot year survey agent, we adopted face-to-face interviews in addition to telephone surveys in the customer experience survey, in order to achieve a broadly representative sample. This was explained in the March 2019 documents.

Evidence from the shadow year survey agent suggests that randomly interviewing a representative group of people by telephone is becoming more challenging because fewer people have landlines, fewer people answer mobile calls from unknown numbers, and interviewing by telephone is less effective at capturing younger
customers. The shadow year survey agent recommends that face-to-face methodology only is used for the youngest age band (18-29).

We acknowledge the difficulties in interviewing people by telephone, however using face-to-face interviews only for any age group can limit the geographical spread of customers interviewed. For this reason we decide to use a mix of telephone and face-to-face interviews for the customer experience survey.

**Our final determination decision**

We decide to use a mix of face-to-face and telephone interviews as a response to the changing way customers are using different communication channels. We will work with the agent appointed to run the C-MeX and D-MeX surveys to determine the appropriate response to this issue.

### 4.3.2 Digital contacts

**What we said in our draft determinations**

Because the way customers contact their company may change over time, we asked the shadow year survey agent to monitor the rates at which digital contacts are replacing non-digital contacts in the shadow year.

**Stakeholders’ representations**

No stakeholders comment on this issue.

**Our assessment and reasons**

The shadow year survey agent advises that the percentage of digital contacts has not decreased in the shadow year to date. The shadow year survey agent also notes that continuing to have the digital quota vary according to sample, and applying its recommendation to align weightings to quotas (which we discuss in section 4.1), should future-proof against changes in the profile of digital and non-digital contacts.

Because we decide to align quotas and weightings, we consider this is an appropriate way to ensure that the customer service survey reflects future changes in the ways customers may contact their company.
Our final determination decision

We do not change our approach, other than to ensure alignment between quotas and weightings for digital customers as we set out in section 4.1.

4.3.3 Representative age bands

What we said in our draft determinations

To ensure the results of the customer experience survey are reflective of each company’s demographic profile we set quotas on the number of interviews by gender and age.

Therefore, we set quotas for the number of interviews by gender and age. We initially based these on three age bands – 18-29 year olds, 30-59 and 60+.

In our draft determinations we said the shadow year survey agent would investigate whether more than three age bands would give more representative results.

Stakeholders’ representations

No stakeholders comment on this issue.

Our assessment and reasons

The shadow year survey agent uses four age bands when conducting the shadow year surveys and says it appears to work well. The shadow year survey agent recommends those four age bands are used:

- 18 to 29;
- 30 to 44;
- 45 to 64; and
- 65 plus.

The shadow year survey agent advises that four age bands enable a more representative sample than three age bands, and that no issues, such as complexity, have arisen from using this approach in the shadow year.
Our final determination decision

We apply the following four age band quotas in the customer experience survey: 18 to 29, 30 to 44, 45 to 64 and 65 plus.

4.4 Higher performance payments

In our draft determinations we said companies can earn standard outperformance payments of up to 6% of allowed residential retail revenues in any one year, based on their performance relative to the other water companies.

We also said higher performance payments of up to 12% will be available to companies that can pass the following gates:

- the company is one of the top three performers by C-MeX score;
- the company is at or above a cross-sector threshold of customer satisfaction performance based on the all-sector upper quartile (ASUQ) of the UK Customer Satisfaction Index (UKCSI); and
- the company has lower than the industry average number of household complaints (per 10,000 connections).

4.4.1 UK Customer Satisfaction Index (UKCSI)

What we said in our draft determinations

On the cross-sector threshold for higher performance payments, we said we would base it on an all-sector upper quartile (ASUQ) from the UK Customer Satisfaction Index (UKCSI) which tracks customer satisfaction across a range of sectors.

We said we would calculate the threshold on an annual basis when applying in-period payments for C-MeX. We said the C-MeX ASUQ will be calculated using the following formula (where SD is standard deviation):

\[
C-MeX\ ASUQ = C-MeX\ mean + (UKCSI\ ASUQ - UKCSI\ mean) / UKCSI\ SD \\
* C-MeX\ SD
\]

Stakeholders’ representations

Three stakeholders (Bristol Water, Northumbrian Water and Dŵr Cymru) comment on the cross-sector threshold for higher performance payments.
Bristol Water states the cross-sector gate for higher performance payments is unlikely to be met and proposes the C-MeX gate is amended so 12% payments apply for the upper quartile of the all-sector UKCSI, 10% for the average of the all-sector UKCSI and 8% for the level of the top utility company in the UKCSI.

Northumbrian Water states we have set an unreasonably high bar for standard and enhanced outperformance payments. The company disagrees with using the upper quartile performance of the UKCSI as a gateway for higher performance payments. The company notes the lack of a clear, published methodology and that the UKCSI includes desirable goods and services as well as utility companies. Northumbrian Water considers the UKCSI mean would be a more appropriate gateway.

Dŵr Cymru states using the UKCSI is inappropriate and rules out any chance of a water company achieving higher payments. The company also claims the sample size of water companies in the UKCSI is too small and that there is an inbuilt bias against water companies compared to other sectors. The company also claims the UKCSI data shows face-to-face interactions score more highly in terms of customer satisfaction compared to other forms of contact, and water companies are disadvantaged because their customers do not typically engage them in this way.

Dŵr Cymru proposes this performance gate is either dropped altogether or adapted to only use UKCSI surveys with customers that had contacted the company.

Our assessment and reasons

All three stakeholders argue to some degree that basing the cross-sector threshold on the UKCSI all-sector upper quartile will result in too high a bar for companies to achieve.

We do not consider this to be the case. Commenting on the approach to using UKCSI, the pilot year survey agent considers it to represent a stretching but attainable target because the highest performing company based on pilot year data was around 2% away from this threshold. Our analysis of pilot year C-MeX scores and the UKCSI survey results published in July 2019 finds that the best performing company would only need to improve its C-MeX score by around 2 to 3% in order to achieve the cross-sector threshold. Because of how we calculate the threshold, this remains the case even when we remove the net promoter score from the C-MeX calculation.

We therefore maintain our position that the UKCSI all-sector upper quartile is an appropriate metric on which to base the cross-sector threshold for higher performance payments in C-MeX. This further encourages companies that are
already high performers in the water sector to meet the levels expected by their customers based on service levels in the wider economy. We see no benefit in revising the threshold by adopting the UKCSI mean (which would be less stretching) or otherwise adapting UKCSI methodology for the water sector.

We also note that any amendments to the methodology, such as the one proposed by Dŵr Cymru, would have cost and resourcing implications, add complexity, and result in a performance threshold not based on such a well-established and readily accessible source as the UKCSI.

For these reasons, we consider it appropriate to retain our approach. Ultimately, the threshold for higher payments is intended to be stretching and only reward companies for achieving high levels of customer service relative to other sectors.

We are aware that the underlying methodology for the UKCSI may change during the 2020-25 period. We will continue to use future iterations of the UKCSI upper quartile, mean and standard deviation for the purposes of C-MeX. However, if the UKCSI methodology moves away from a league table approach such that we cannot quantify an upper quartile or it no longer has a sufficient number of water companies in its sample, we will use the last appropriate UKCSI results instead in our in-period determinations.

**Our final determination decision**

We maintain our decision that one of the criterion for the three gates for higher performance payments in C-MeX is that a company is at or above a cross-sector threshold of customer satisfaction performance, based on the all-sector upper quartile (ASUQ) of the UK Customer Satisfaction Index (UKCSI).

We will continue to calculate the cross-sector threshold as follows (where SD is standard deviation):

\[
C-MeX\ ASUQ = C-MeX\ mean + (UKCSI\ ASUQ\ -\ UKCSI\ mean) / UKCSI\ SD \\
* C-MeX\ SD
\]

**4.4.2 Customer complaints**

**What we said in our draft determinations**

In our PR19 methodology we said we would not include the number of customer complaints received by a company when calculating its C-MeX score. We were
concerned this may reduce the incentive for companies to proactively contact their
customers or to innovate. Instead, we said companies could earn higher
performance payments if they performed satisfactorily on complaints.

We reiterated this in our draft determinations and confirmed that companies would
only be eligible for higher performance payments if they receive below the industry
average for the number of complaints per 10,000 connections as reported to the
Consumer Council for Water.

**Stakeholders’ representations**

The Consumer Council for Water states it is concerned by the insufficient weight
given to complaints performance within C-MeX. It states companies may be able to
achieve outperformance payments even when their complaints performance is static
or declining.

Northumbrian Water proposes the complaints gateway for higher performance
payments should be based on absolute levels of performance because the company
claims our proposed approach favours companies that have improved from a poor
base.

**Our assessment and reasons**

We decide to include companies’ relative complaints performance in C-MeX as one
of the gates for higher performance payments. This means that companies cannot
receive higher performance payments unless they have less than the average
number of complaints relative to other companies.

As we said in our draft determinations, we remain concerned that including
complaints performance as a specific metric in the main C-MeX score could act as a
perverse incentives by deterring companies from proactively carrying out
engagement activities and initiatives that would benefit customers but may generate
contacts or complaints.

While complaints performance is not a specific metric in the main C-MeX score, we
recognise that a customer’s satisfaction with their company is likely to reflect their
complaints experience, particularly for those who have contacted their company. We
consider this means that while possible, it is unlikely that companies with poor
performance in relation to complaints will receive outperformance payments. We
expect this to be a low risk following the Consumer Council for Water’s recent decision to revise its definition of a complaint.\(^3\)

We note that the total number of complaints will continue to be monitored and published annually by the Consumer Council for Water in comparative league tables.

For these reasons, we maintain the approach in the draft determinations which we consider is an appropriate way to take account of complaints in C-MeX. C-MeX is only one tool to incentivise companies to address customers’ complaints. We use other regulatory tools to incentivise companies to address customers’ complaints:

- Our past delivery assessment in our initial assessment of business plans led us to set action plans for companies to improve (see ‘Accounting for past delivery technical appendix’).
- We set bespoke performance commitments for companies where we identify an issue.
- As we set out in our new strategy for beyond 2020 we commit to working with other regulators and the UK and Welsh Governments to develop long-term targets for improving performance in relation to resilience, the environment and customer excellence. As part of this, we will increase our wider company monitoring and insights approach and consider how to develop it further over the coming years. This includes exploring how we can better utilise complaints data in order to further our understanding of company performance.

Regarding Northumbrian Water’s proposal for how we use complaints for the higher payments gates, we do not base it on a company’s own improvement. Instead a company must be above the sector average as measured by total complaints per 10,000 connections. This ensures companies do not receive higher performance payments for below average performance in complaints compared to the rest of the sector, even if their own performance in relation to complaints has improved.

**Our final determination decision**

We maintain our decision to include companies’ complaints performance in C-MeX as one of the gates for higher performance payments. This is measured as being below the industry average for the number of household complaints that each company receives as measured by the Consumer Council for Water per 10,000 connections.

---

\(^3\) See ‘Complaint reporting guidance’, the Consumer Council for Water (September 2019)
4.4.3 Final decision for higher performance payments

As we set out above in sections 4.4.2 and 4.4.3, we maintain our decision for the following gates for higher performance payments in C-MeX:

- the company is one of the top three performers by C-MeX score;
- the company is at or above a cross-sector threshold of customer satisfaction performance based on the all-sector upper quartile (ASUQ) of the UK Customer Satisfaction Index (UKCSI); and
- the company has lower than the industry average number of household complaints (per 10,000 connections).

As a consequence of revising our approach to applying financial incentives for C-MeX (see section 3.1) we make minor revisions to how we proposed, in our draft determinations, to award higher performance payments.

Because of the stepped approach we proposed in our draft determinations, companies that had not achieved 6% in standard outperformance payments would have achieved up to 3% (the next step down). To apply higher performance payments of 8%, 10% or 12% would have been more than doubling the performance payments for those companies, and for that reason we said the maximum they could earn would be 6% (doubling their standard outperformance payments) if they met all three gateways. Because we no longer apply the stepped approach, we do not need to limit higher performance payments in this way.

As a result, under our revised approach the company with the highest score that passes the three gates receives an additional 6% of that year’s annual allowed residential retail revenue, potentially taking its total outperformance payments to 12%. If a second company qualifies, it will receive an additional 4% and if a third company qualifies it will receive an additional 2%. For the avoidance of doubt, if only one company passes the three gates it will receive an additional 6% on top of its standard outperformance payment, regardless of whether it is has the highest C-MeX score across all companies.
5. **Design decisions for D-MeX**

In this section we consider stakeholder representations and design issues that are specific to D-MeX.

D-MeX measures the experience of customers of developer services provided by the 17 largest water companies in England and Wales. These customers include small and large property developers, self-lay providers, new appointees and some residential customers that have new mains connections installed.

A self-lay provider is a contractor that is appointed to carry out on-site work on behalf of a developer. A new appointee is a small company that can also be appointed to carry out on-site work and subsequently becomes the monopoly supplier for that development as a result of a variation of the incumbent monopoly provider’s supply area. Both self-lay providers and new appointees transact with the incumbent when connecting to its network, and new appointees normally arrange a bulk supply or discharge agreement with the incumbent to supply the site.

5.1 **Qualitative component – the customer satisfaction survey**

5.1.1 **Representation of larger customers**

*What we said in our draft determinations*

As part of conducting the satisfaction survey for D-MeX, the shadow year survey agent receives the contact details for customers who have transacted with a water company in the previous month. The shadow year survey agent removes from the sample the names of contacts that have been surveyed more than once in the last six months, whether in relation to the same or a different water company (this process is referred to as ‘de-duplication’). The purpose of undertaking de-duplication is so no customer is surveyed more than once in six months.

However some customers may have transacted with more than one water company in the six month period – this is most likely to apply to larger developers, new appointees and self-lay providers. The approach to de-duplication described above means those customers can only be in the sample once in a six month period, no matter how many companies they may have dealt with. This means that when they are selected for an interview about one company, they will be unable to provide a view on other companies during this period.
In our draft determinations, we recognised this could lead to an under-representation of these customers in the survey results. We said one option to address this could be to ask about more than one transaction or more than one company for these customers when they are surveyed; or to allow these customers to be in the sample more than once in six months.

We asked the shadow year survey agent to advise us on possible approaches to this issue and asked for feedback from stakeholders.

**Stakeholders’ representations**

Independent Water Networks, a new appointee, supports lifting the restriction of a customer only appearing a maximum of once in six months in a company’s survey sample. The respondent considers that providing feedback on multiple companies or transactions in a single phone call could be confusing and would not save time.

Bristol Water states a multi-transaction survey would be impractical and unworkable.

Fair Water Connections states our survey approach heavily skews against work undertaken by self-lay providers and new appointees. The Consumer Council for Water questions how D-MeX would reflect the experience of self-lay providers and new appointees.

**Our assessment and reasons**

The shadow year survey agent considers that surveying customers about multiple transactions would require a separate study to be conducted. Although this could mitigate the potential for under-representation, doing so would introduce complexities to calculating the D-MeX score.

The shadow year survey agent recommends the six-month exclusion period for repeat surveying customers that have dealt with more than one company during this time is reduced to two months. This would mean developer services customers would only be in the sample once in six months for any particular water company but could be in the sample for up to three water companies during that six-month period. The shadow year survey agent considers an exclusion period less than two months is not best survey practice because it may overburden respondents and could lead to bias in responses due to the influence of previous interviews.

We agree with this approach and consider it addresses the issue in a proportionate and practical way. We also agree with the shadow year survey agent that this should apply to all developer services customers. This change will increase the chance of
being selected for surveys for developer services customers who deal with more than one water company, and potentially allow them to provide views on up to three companies during a six-month period.

In addition, the representation of self-lay providers and new appointees in D-MeX as a whole is likely to increase because we include new metrics for self-lay providers in the data source for the qualitative element of D-MeX, and expect to include new metrics for new appointees in the future (see section 5.1.4).

**Our final determination decision**

We reduce the repeat survey exclusion period for all developer services customers from six to two months, in relation to any water companies about which they have not been interviewed in the last six months. We maintain a repeat survey exclusion period of six months in relation to a customer and the water company about which they were interviewed.

**5.1.2 Weighting survey responses**

**What we said in our draft determinations**

In our draft determinations we considered whether the views of larger customers responsible for multiple connections per development may be under-represented in the qualitative survey results, and whether their views should carry more weight in the qualitative survey.

In our draft determinations we decided not to apply weightings to address this issue. We identified significant variation across companies in the size of their customers which would have suggested using different weightings for each company; this would have undermined our non-company specific approach for D-MeX. We also wanted to ensure all customers are treated as equally important, regardless of their size or level of business with a company, so all customers receive excellent service.

We asked the shadow year survey agent to advise us on potential approaches to this issue.

**Stakeholders’ representations**

Fair Water Connections expresses concerns that D-MeX will struggle to match both surveys and metrics to the value of work each customer category receives or handles.
Thames Water proposes a relationship survey which would ask customers to consider their overall relationship with the company in question instead of being asked about their experience based on service delivery. The company considers this is necessary because the transactions survey does not give adequate voice to some customer groups including self-lay providers and new appointees.

**Our assessment and reasons**

We want companies to treat all developer services customers as important, regardless of their customer type or size. We have heard from some companies that, in the normal course of business, they foster strong commercial relationships with larger developer customers. We also want to avoid larger customers having a disproportionate impact on survey results, which would reduce the voice of small customers for which very poor service is likely to have a greater impact. A higher weighting for larger customers risks encouraging companies to focus on larger customers which may come at the cost of service to smaller customers.

The shadow year survey agent recommends not applying weighting based on customer size because defining large, medium and small customers across companies has not been feasible. The shadow year survey agent also states that weighting customers by volume of transactions with a company, which it considers would be the most objective way to weight by customer size, could mean a very small percentage of interviews would carry a disproportionate weight of the survey results.

We consider that other features of D-MeX (such as reducing the repeat survey exclusion period to increase the chances of being selected in surveys; and maintaining an equal weighting for the quantitative component of D-MeX) will go some way to address concerns that larger customers may be under-represented in D-MeX. Because larger customers are more likely to undertake multiple transactions with any one company, this will be reflected in the calculation of the quantitative component of D-MeX which uses all data submitted, rather than a sample.

In response to Thames Water’s proposal, we do not consider a relationships survey to be appropriate for D-MeX – the qualitative component is designed to measure the quality of customer service provided rather than the quality of a company’s overall relationships with its developer services customers.
Our final determination decision

For the reasons described above we maintain our position at draft determinations. We will not apply weightings to customer responses in the developer customer satisfaction survey.

5.1.3 Sample size in the qualitative survey

What we said in our draft determinations

In our draft determinations we said that subject to confirming estimates for the confidence intervals for each company during the shadow year, and considering what is practical to achieve, we were minded to maintain our approach of surveying 20% of a company’s customers, boosted to 100 customers per year where 20% falls below this number.

We also asked the shadow year survey agent to recommend an appropriate sample size (or quota) for D-MeX to apply to the 2020-25 period.

Stakeholders’ representations

Bristol Water states this sample size may not be achievable and that asking the same developer customers repeatedly might drive satisfaction ratings down.

Our assessment and reasons

The shadow year survey recommends a target sample size of 20% of usable contacts except for two cases. For the bigger companies, 20% of sample is more than 1,000 interviews per year above which there are diminishing returns. However, there is an issue with some small companies where surveying 20% of their customers could in some cases lead to monthly samples of fewer than 10 interviews. This leads to very small effective sample sizes, which reduce confidence in the survey results for smaller companies.

The shadow year survey agent proposes increasing the percentage of contacts provided to be included in the sample for the three smallest companies. The shadow year survey agent proposes this is increased to at least 30% of the sample with the aim of achieving as many interviews as possible from the contacts available.

The sample will be based on 20% of those customers that transact with the company in a month. For each company, where 20% of the sample is expected to be more
than 1,000 in a year, the sample will be capped at 1,000; where 20% of the sample is expected to be less than 100 in a year, the sample will be increased to at least 100 in a year.

We consider this revised approach will make the sample size achievable, and improve accuracy of D-MeX scores for the smaller companies. In our decision to cap the sample size for the larger companies, we seek to balance the costs of these surveys and the accuracy of results, noting the shadow year survey agent considers that a sample size above 1,000 per year would deliver diminishing returns.

In terms of whether this approach may affect satisfaction scores, our approach to the exclusion period means that developer services customers will not be contacted more than twice a year about the same company. This reduces the risk of the contact frequency affecting their satisfaction ratings.

**Our final determination decision**

Following our assessment, the sample will be based on 20% of a company’s monthly developer services customers. For each company, where 20% of the sample is expected to be more than 1,000 in a year, the sample will be capped at 1,000; where 20% of the sample is expected to be less than 100 in a year, the sample will be increased to at least 100 in a year.

**5.1.4 Data source for the qualitative survey**

**What we said in our draft determinations**

Since 2015, water companies in England and Wales have voluntarily reported on the services they provide to developer customers. Companies report against a set of metrics developed by water companies, with public reporting administered by Water UK. The metrics measure the compliance of water companies with levels of service for a range of developer service activities.

Each month companies provide the relevant survey agent with a list of contact details for customers who have completed transactions with the company in the previous month. A set of these metrics will be the ‘transactions’ for the purposes of the qualitative survey in D-MeX – the survey agent uses these contacts as the source of the sample for the survey.

In our *March 2019 guidance document* we set out that for the shadow year we would use all performance-based metrics excluding those relating to acknowledgements.
For completeness we also included a further four metrics relating to self-lay providers.

**Stakeholders' representations**

No stakeholders comment on this issue.

**Our assessment and reasons**

We consider it appropriate to include all Water UK metrics that are performance-related excluding those relating to acknowledgements as the data source for the qualitative component of D-MeX.

While this maintains the approach we used in the shadow year, it is not necessarily the same as our approach for the quantitative component (discussed below). This is because the survey focuses on the quality of customer service provided to developer customers rather than just the physical service delivery; it is therefore appropriate that it relates to a broader set of metrics.

**Our final determination decision**

For the data source for the qualitative element of D-MeX, we include all current performance-based Water UK metrics, excluding most acknowledgements. We set these out in annex 1 to this document.

In terms of new metrics that may be introduced by Water UK in the 2020-25 period, we use the same revisions process and criteria as described in section 5.2.1 for the quantitative metrics, to assess any proposed changes.

The shadow year survey agent has recommended that the date companies are required to provide the monthly sample be brought forward. This issue will be dealt with in guidance, to be finalised by March 2020.
5.2 Quantitative component – the levels of service metrics

5.2.1 Review of the Water UK levels of service metrics

What we said in our draft determinations

In our PR19 methodology we said we would use company performance against some or all of the metrics for the quantitative component of D-MeX.

In our March 2019 documents we said we will carry out a review of the metrics prior to their inclusion in D-MeX to ensure they are fit for purpose. We engaged Jacobs to undertake this review. We also said we would use all existing performance-based metrics as the basis for the quantitative component of D-MeX in the shadow year and future iterations of them. In our draft determinations we said the metrics that will go into D-MeX will be the metrics in place at the time our final determinations in December 2019, and that the review we propose to undertake would inform our final decision on the way the Water UK metrics are used in the D-MeX calculation.

Stakeholders’ representations

Fair Water Connections expresses concern that the current metrics do not match the value of work each customer category receives and handles, and that they appear weighted to relatively easy activities, such as providing single connections.

Our assessment and reasons

In this section we consider:

- which metrics should be included in the quantitative component of D-MeX, following our technical consultant’s review of the Water UK metrics;
- our approach to managing changes to the metrics during the 2020-25 period; and
- our approach to managing different statutory requirements for Welsh and English companies that affect the metrics.

In its review of the Water UK metrics, Jacobs recommends which metrics to include for the purposes of D-MeX. We commissioned this work in order to inform our final decision on weighting in the qualitative element and weighting between the qualitative and quantitative elements of D-MeX taking into account the overarching principles for C-MeX and D-MeX.
To inform its assessment, Jacobs consulted with developer services customers and water companies through online surveys, one-to-one interviews and our D-MeX working group. Jacobs also reviewed water companies’ past performance as reported to Water UK. In its final report, which we publish alongside our final determinations, Jacobs recommends we include 26 metrics in D-MeX, of which six are draft metrics relating to services provided to new appointees.

The set of metrics proposed for inclusion in D-MeX represents the main elements of service delivery across the different customer types. In general, metrics were not recommended for inclusion if they measured activities that were seen as low effort by companies or low importance to customers — such as those focused on sending acknowledgements.

To ensure coverage and customer representation the Jacobs report includes a recommendation on new metrics for new appointees, which were developed by Water UK and provided in draft to Jacobs in August 2019. However, we understand these proposed metrics for new appointees will not be formally approved by Water UK until 2020, and as a result they are not included in D-MeX at this time.

Because the recommended metrics sufficiently represent the main elements of service delivery and exclude those seen as low effort by companies or low importance to customers, we consider the recommendations made by Jacobs to be appropriate and we accept them. We therefore include in the quantitative component of D-MeX all metrics recommended by Jacobs, with the exception of the proposed metrics for new appointees, for the reasons given above.

We may need to manage changes during the 2020-25 period. The metrics may be revised by Water UK, for example the proposed new metrics relating to new appointees. Changes to metrics should reflect rising customer expectations and service levels, which in turn should be reflected in D-MeX. As a result, we consider revisions or additions to the metrics in D-MeX may need to be incorporated in D-MeX during the 2020-25 period.

In determining whether to make changes to D-MeX as a result of changes to the Water UK metrics, our decision will be based on the principles that revisions are in customer interests, support consistent and fair comparisons between companies and align with our wider duties. We will consult with companies proportionately, taking into account previous stakeholder engagement. See annex 2 of the ‘Outcomes performance commitment appendix’ for each company for more information on how we will manage changes to performance commitments in the 2020-25 period.
We next consider the implication of different statutory requirements between companies operating wholly or mainly in Wales (‘Welsh companies’) and companies operating wholly or mainly in England (‘English companies’).

An example of upcoming changes that will result in different requirements for Welsh and English companies are new or revised metrics arising from the Code for Adoption Agreements for English companies. On 25 October 2019 we approved sector guidance and a model adoption agreement for wastewater services provided by self-lay providers. These documents and metrics were developed by Water UK in consultation with self-lay providers and their representatives. We expect to progress sector documents for water services under the Code for Adoption Agreements later in 2020.

We recognise that the Code for Adoption Agreements for English companies does not apply to Welsh companies. We therefore decide that any revisions to existing or new metrics introduced as a result of this code will not apply for Welsh companies for the purposes of D-MeX. To avoid adverse impacts on the service provided to their customers, Welsh companies should instead report against the equivalent metrics, prior to the changes being made.

Should new metrics be developed as a result of the introduction of a Code for Adoption Agreements for Welsh companies, we expect to include them in D-MeX. While this could be seen to reduce the comparability between Welsh and English companies, we do not consider it appropriate for Welsh companies to be measured by service levels developed under the English Code for Adoption Agreements. Instead, Welsh and English companies will in effect be compared on the basis of their compliance with metrics aligned to the standards relevant to them. We consider this further in section 5.3.2.

We will use our approach to managing changes to the metrics during the 2020-25 period as set out above to confirm which metrics apply for English and Welsh companies where there is a difference.

Our final determination decision

We include the metrics recommended by Jacobs in the quantitative component of D-MeX, with the exception of the proposed metrics for new appointees, for the reasons given above.

We set out the metrics in the quantitative component of D-MeX that will apply to all companies from 1 April 2020 in annex 2 of this document.
In determining whether to make changes to D-MeX as a performance commitment in light of changes to the Water UK metrics, our decision will be based on the principles that revisions are in customer interests, support consistent and fair comparisons between companies and align with our wider duties. We will consult prior to making any changes to D-MeX.

### 5.2.2 Calculating the quantitative component of D-MeX

#### What we said in our draft determinations

In our March 2019 documents we said the quantitative component of D-MeX will be calculated using a simple average of the Water UK metrics.

In our draft determinations we responded to stakeholder concerns that this approach could underrepresent activities that are important to customers or underrepresent types of customers such as self-lay providers or new appointees. We noted problems with adopting alternative approaches and said we were minded to use an ‘average of averages’ approach, where different metrics are grouped into categories and an average is taken of these categories.

We said our final decision would be subject to our review of the metrics during the shadow year and stakeholder feedback to our draft determinations.

#### Stakeholders’ representations

Bristol Water proposes using the overall score (that is, a simple average) while United Utilities supports an ‘average of averages’ approach.

#### Our assessment and reasons

An average of averages approach attempts to increase the impact of certain metrics, to reflect the importance of the underlying activity to customers or ensure adequate representation of customer types. The set of metrics we decide to include in D-MeX are selected based on their importance to customers and other criteria as set out in section 5.2.1. Accordingly we see less added benefit from an average of averages or weighted approach.

In its review, Jacobs considers aggregating metrics into combined service areas but on balance does not recommend combining metrics due to the additional level of complexity and reduction in transparency this would introduce.
Our final determination decision

We apply a simple average approach for calculating the quantitative component of D-MeX.

5.3 Incentive design

5.3.1 Weighting between D-MeX components

What we said in our draft determinations

In our March 2019 documents we said that subject to results of our review of the Water UK metrics we were minded to apply an equal 50:50 weighting for both the quantitative and qualitative components of D-MeX.

In our draft determinations we said our final decision will be subject to both feedback on our draft determination and the results of our review of the metrics.

Stakeholders’ representations

Fair Water Connections states the Water UK metrics are not sufficiently demanding because companies score highly on them.

The Consumer Council for Water asks how we will ensure companies are incentivised to perform as well as possible and not just ‘sufficiently’ well in order to meet the minimum requirement.

Our assessment and reasons

We consider the relative weightings between the quantitative and qualitative components of D-MeX based on their contribution to overall stretch and the relative representation of customer types.

While high scores reported by companies in recent years against the Water UK metrics indicate high levels of service, they may also imply the metrics are insufficiently stretching or do not adequately reflect comparative performance.

In its review of the Water UK metrics, Jacobs notes that in terms of the overall level of stretch, the metrics have not been subject to external assurance processes and some target levels are likely to be tightened by changes following the English Code
for Adoption Agreements. Jacobs concludes that overall, the set of metrics provides sufficient stretch at this point in time and there is no need to reduce the weight of the quantitative component of D-MeX.

We also consider the effect on representation of customer types. Some stakeholders argue our sampling approach to the qualitative survey reduces the representation of larger customers such as self-lay providers, new appointees and large developers. The quantitative metrics ensure that those customers are better represented because the quantitative component is based on the volume of transactions and includes specific metrics for those types of customers.

Our final determination decision

Following our assessment, we maintain our decision as set out in our draft determinations. We will apply a 50:50 weighting for the quantitative and qualitative components of D-MeX.

5.3.2 Comparability of Welsh and English companies

What we said in our draft determinations

In our PR19 methodology, we said we were not minded to apply company-specific or other adjustments to D-MeX to avoid undue complexity.

In our March 2019 document, we said we were not minded to treat companies differently but would continue to monitor the impact of the D-MeX survey on companies in the shadow year before making a final decision.

In our draft determinations, we said we found little evidence to suggest that legal and regulatory differences between Wales and England put either Welsh or English companies at a disadvantage in relation to D-MeX. We said we were minded to treat Welsh and English companies in the same way and include all companies in a single D-MeX league table.

Stakeholders’ representations

Dŵr Cymru states Welsh companies cannot be fairly compared with English companies in D-MeX. The company claims the survey responses of its developer services customers during the shadow year show negative reactions by customers to the legal and regulatory requirements in Wales, such as the arrangements for sewer adoptions, domestic sprinklers and surface water management.
Our assessment and reasons

We first consider the comparability of Welsh and English companies based on results from the qualitative component of D-MeX (the survey) in the shadow year. We then consider comparability in the context of the quantitative component (the Water UK metrics) drawing on our decisions in section 5.2.2.

In terms of the qualitative survey, we consider the evidence presented to us by Dŵr Cymru in its representation. We also ask the shadow year survey agent to review shadow year data to assess whether the regulatory requirements in Wales and England have a noticeable impact on D-MeX scores.

Dŵr Cymru includes comments from developer services customers in its representation, indicating frustration with the legal or regulatory requirements that affect Welsh companies. However in principle we do not consider this leads to unavoidable disadvantage for Welsh companies – primarily because companies are able to manage potential customer dissatisfaction through effective communication and good customer service.

We also note that upcoming changes for English companies relating to new connections charging rules and adoption agreements may cause increased customer dissatisfaction for English companies, which similarly we consider is well within company control to manage.

In addition, from its review of responses to the shadow year, the shadow year survey agent finds no evidence to suggest that the different regulatory frameworks in Wales and England impact D-MeX scores in such a manner that would require calibration or some other intervention.

On this basis, for the qualitative survey, we find little evidence of issues caused by the legal and regulatory differences between Wales and England that are not within company control to manage.

We discuss reporting against the quantitative metrics, for Welsh and English companies in section 5.2.1 above.

It could be argued that this approach to the quantitative metrics reduces the comparability between English and Welsh companies, particularly if the metrics Welsh companies report against are seen to be less stretching than the revised metrics English companies report against. We recognise this risk but do not consider it appropriate to require Welsh companies to be measured against service levels developed under the Code for Adoption Agreements for English companies. Should
a Welsh Code for Adoption Agreements come into force during the 2020-25 period,
Welsh companies will report against standards relevant to them. This approach
means that in effect Welsh and English companies will be compared on the basis of
their compliance with their respective statutory requirements.

On this basis, we consider our approach to the quantitative metrics does not overall
diminish the comparability of Welsh and English companies in D-MeX.

In summary based on our consideration of both components of D-MeX, we consider
any differences in legal or regulatory arrangements for Welsh and English
companies is well within management control in terms of the qualitative survey, and
the nature of comparisons and extent of difference in the quantitative metrics does
not overall reduce comparability between Welsh and English companies.

For these reasons we consider it appropriate to maintain a single league table for
Welsh and English companies. We note the significant benefits created by a single
league table for Welsh and English companies through strong incentives on
companies to improve their performance relative to all other companies.

We consider alternative approaches such as Dŵr Cymru’s proposals in its December
2018 letter to us to set company-specific targets or exclude Welsh companies from
the league table do not provide sufficient incentives, and would weaken incentives
on all companies to improve by removing relative benchmarks.

Our final determination decision

We maintain our decision as set out in our draft determinations. We include both
Welsh and English companies in a single league table.

5.3.3 Size of financial incentives

What we said in our draft determinations

In our PR19 methodology we said D-MeX outperformance payments would be up to
2.5% and underperformance payments would be up to 5% of a company’s annual
developer services revenue.

In our March 2019 documents we said there was a good case to increase the size of
financial incentives to reflect the high impact that poor service to developers can
have on the wider economy and society, in particular housing supply. In our draft
determinations we proposed D-MeX outperformance payments should be up to 6% and underperformance payments up to 12% of annual developer services revenue.

**Stakeholders’ representations**

Three stakeholders express concerns with our proposed size of financial incentive:

- Bristol Water states the rationale for change has not been set up clearly and notes the lack of timely signalling of this change;
- Dŵr Cymru and South West Water state that we should not increase the size of financial incentives because D-MeX is a new mechanism; and
- South West Water proposes we set more balanced incentive rates between outperformance and underperformance rates.

**Our assessment and reasons**

We assess the appropriate size of D-MeX financial incentives based on whether it provides a material incentive to change company behaviours and drive meaningful improvements in customer priority areas, including whether the size of the financial incentive is proportionate in the context of companies' overall incentive packages, while reflecting the importance of housebuilding to society and the wider economy.

Incentive levels at +6% to -12% of developer services revenue could see maximum payments for companies of +0.1% to -0.2% return on regulatory equity each year, which is a modest component of companies’ overall incentive packages. We consider that reducing the size of incentives to the previous position would materially weaken incentives on companies to improve their performance. In particular doing so would undermine our objective to incentivise service improvements for developer services customers, having regard to the UK Government’s strategic priorities relating to support for housebuilding in England.

While one stakeholder challenges how we communicated the change, we did signal a possible increase in our March 2019 documents and invited comments on our proposal in our draft determinations. In response, three stakeholders object to the increase but do not provide sufficient and convincing evidence for why a smaller incentive would be more appropriate.

Two stakeholders comment that D-MeX is a new performance commitment. We consider that the components of D-MeX have now been well-tested after pilot and shadow year surveys. The quantitative metrics have been reported by companies since April 2015 and have been reviewed and updated by industry since then. Taken together we consider the overall D-MeX measure, while new, is sufficiently robust,
and has been sufficiently communicated and tested, to justify the proposed size of incentives.

In terms of the balance between underperformance and outperformance payments, in our PR19 methodology we said we would apply asymmetric incentives with greater underperformance than outperformance payments:

- first, because this was a design feature of the SIM that has proven effective at incentivising improvements in performance; and
- secondly, because of the poor service received by developer services customers.

For these reasons we maintain our position that asymmetric incentives are appropriate, although as we set out in section 3.1 on how we will apply incentive payments, our revised approach means the asymmetry principle relates to the size of incentive payments rather than the distribution between companies. This provides a good incentive for companies to aim to avoid underperformance payments from ranking below other companies.

**Our final determination decision**

Following our assessment, we maintain our decision as set out in our draft determinations. Outperformance and underperformance payments for D-MeX are between +6% and -12% of a company’s annual developer services revenue.

**5.3.4 Developer services revenue**

**What we said in our draft determinations**

In our PR19 methodology we said we would base outperformance and underperformance payments on a proportion of a company’s developer services revenue, which is what a company collects from developer services customers each year for the delivery of contestable and non-contestable developer services.

In our March 2019 documents we confirmed the definition of developer services revenue for the purposes of D-MeX as the sum of water and wastewater connection charges, infrastructure charge receipts, charges for requisitioned mains and sewers, diversions and other price control contributions.

Some stakeholders raised concerns that companies with a high level of self-lay provider activity, and therefore lower developer services revenue from contestable services, may as a result receive lower outperformance or underperformance
payments. In our draft determinations we said we were minded to continue to base financial incentives on annual developer services revenue.

**Stakeholders’ representations**

Bristol Water supports our approach to base financial incentives on developer services revenue, but asks for clarity on how it will work in practice particularly with our revised approach to the treatment of developer services in the price controls.

Consumer Council for Water and Fair Water Connections note that under our approach companies will receive lower outperformance and underperformance payments if there is a higher level of competitive developer services in its supply area. Consumer Council for Water asks how D-MeX will incentivise the right behaviour by companies.

**Our assessment and reasons**

We do not consider there is a practical alternative to using a company’s annual developer services revenue as the basis for incentive payments, particularly to reflect the level of developer services activity across companies. No alternatives were proposed in representations.

Even though companies may receive lower outperformance or underperformance payments as a result of higher levels of contestable activity in their supply area, we consider other design features of D-MeX such as the reputational league table will still provide incentives on companies to improve their service performance.

As we said in our draft determinations, we consider that any risk to competition caused by our use of developer services revenue to calculate D-MeX incentive payments is reduced by our wider policies and other regulatory tools in place which enable competition in new connections and ensure a level playing field for new entrants. Because of this, our choice of developer services revenue for D-MeX payments should not drive perverse behaviours by companies with relation to competition and service delivery.

Regarding our decisions relating to the treatment of developer services in the price controls (see ‘Our approach to regulating developer services’) we note that:

- We base incentive payments for D-MeX on actual developer services revenue collected each year by companies. There is therefore no interaction with the revenue allowance we make in our final determinations and the developer services end-of-period reconciliation that will occur at PR24.
• Developer services revenue will no longer include income from non-section 185 diversions. We do not consider it necessary to include this in developer services revenue for the purposes of D-MeX because the predominantly statutory users of these services are not included in the qualitative or quantitative components of D-MeX and it would add unnecessary complexity.

**Our final determination decision**

Following our assessment, we maintain our decision as set out in our draft determinations. We calculate incentive payments as the relevant proportion of a company’s annual actual developer services revenue as reporting for the following components (for each of water and wastewater):

- connection charges;
- infrastructure charge receipts – new connections;
- requisitioned mains;
- requisitioned sewers;
- diversions; and
- other contributions (price control).

As set out above, the definition of diversions is aligned to what is included in the PR19 price controls.

For the avoidance of doubt, developer services revenue should be reported gross of any discounts or income offsets for the purposes of D-MeX.

**5.3.5 Reconciling financial incentives**

**What we said in our draft determinations**

In our draft determinations we said that we would apply an aggregate sharing mechanism for all outcome delivery incentives, to protect customers from higher than expected outperformance payments.

We said we would not include performance commitments at the retail price control level in the sharing mechanism but all other performance commitments are in scope. However, as we set out in our PR19 methodology, outperformance and underperformance payments for D-MeX are funded through the water network plus and wastewater network plus controls.
Companies proposed price control allocations for D-MeX in their September 2018 business plans, and these ranged from a 23-77 to 64-36 water-wastewater split. Hafren Dyfrdwy allocated D-MeX to its business retail control.

**Stakeholders’ representations**

No stakeholders comment on this issue with relation to D-MeX.

**Our assessment and reasons**

We consider customers to be protected from higher than expected outperformance payments for D-MeX because outperformance payments are capped at 6% of developer services revenue and the relative impact of potential D-MeX in companies’ overall incentive packages is small. It would add complexity to include D-MeX in the aggregate sharing mechanism for little additional customer benefit.

We consider Hafren Dyfrdwy should not have allocated its D-MeX performance commitment to business retail control because we said in our PR19 methodology this should be allocated to the water network plus and wastewater network plus controls. We decide to change the price control allocation for this company to water network plus and wastewater network plus.

We reflect further on the appropriate price control allocations for D-MeX given we operate D-MeX as an in-period incentive based on how companies perform relative to other companies and with outperformance and underperformance payments as a proportion of actual developer services revenue collected each year. We note that developer services are volatile due to the contestability and uncertainty around new housing developments and are subject to considerable forecasting risk.

Because customers bear the outperformance and underperformance payments that each company receives or incurs, we consider it necessary to more closely align the payments to the price control that applies to those customers. As part of in-period reconciliation we therefore allocate them between the water network plus and wastewater network plus controls based on the outturn revenues (as already reported by companies in their annual performance reports), rather than the price control allocations as submitted by companies in their September 2018 business plans.

**Our final determination decision**

As part of in-period reconciliation for D-MeX:
• we do not include D-MeX outperformance payments in the aggregate sharing mechanism for outcome delivery incentives; and
• we vary price control allocations for all companies based on the proportion of developer services revenues collected by companies for either water or wastewater services each year.
6. **Next steps**

We will publish the guidance by March 2020. The guidance will support companies to provide the necessary information to achieve an accurate and representative set of C-MeX and D-MeX results in the 2020-25 period.

Both C-MeX and D-MeX will come into effect on 1 April 2020 and operate until 31 March 2025. As performance commitments with in-period financial incentives, we will need to reconcile the outperformance or underperformance payments received or incurred by companies during the 2020-25 period.

Alongside other PR19 mechanisms, we will consult on the models and our approach to reconciliation in early 2020.
Annex 1: Qualitative metrics in D-MeX

As discussed in section 5.1.4, these metrics are the data source for the qualitative component of D-MeX from 1 April 2020.

- W1.1 Pre-development enquiry – reports issued within target
- W3.1 s45 quotations – within target
- W4.1 s45 service pipe connections – within target
- W6.1 Mains design <500 plots – quotations within target
- W7.1 Mains design >500 plots – quotations within target
- W8.1 Mains construction within target
- W17.1 Mains diversions (without constraints) – quotations within target
- W17.2 Mains diversions (with constraints) – quotations within target
- W18.1 Mains diversions - construction/commissioning within target
- W20.1 Self-lay Point of Connection report < 500 plots etc - reports issued within target
- W21.1 Self-lay Point of Connection reports >500 plots etc - reports issued within target
- W23.1 Self-lay design and terms request <500 plots etc – quotations within target
- W24.1 Self-lay design and terms request >500 plots etc – quotations within target
- W25.1 Self-lay signed agreement – acknowledgements within target
- W26.1 Self-lay water for pressure/bacteriological testing – provided within target
- W27.1 Self-lay permanent water supply – provided within target
- W28.1 Self-lay vesting certificates – issued within target
- W29.1 Self-lay Asset Payments – issued within target
- W30.1 Self-lay plot references and costing details – issued within target
- S1.1 Pre-development enquiry – reports issued within target
- S3.1 Sewer requisition design – offers issued within target
- S4.1 Sewer requisition – constructed and commissioned within agreed extension
- S6.1 Technical vetting of adoptions & diversions – approval or rejection letters within target
- S7.1 Adoption legal agreement – draft agreements issued within target
- S8.1 s106 sewer connection – approval letters issued within target
- S9.1 s106 sewer connection – rejection letters issued within target
Annex 2: Quantitative metrics in D-MeX

As discussed in section 5.2.1, these metrics form the basis of the **quantitative component** of D-MeX from 1 April 2020.

- W1.1 Pre-development enquiry – reports issued within target
- W3.1 s45 quotations – within target
- W4.1 s45 service pipe connections – within target
- W6.1 Mains design <500 plots – quotations within target
- W7.1 Mains design >500 plots – quotations within target
- W8.1 Mains construction within target
- W17.1 Mains diversions (without constraints) – quotations within target
- W17.2 Mains diversions (with constraints) – quotations within target
- W18.1 Mains diversions – construction/commissioning within target
- W20.1 Self lay Point of Connection report < 500 plots etc – reports issued within target
- W21.1 Self lay Point of Connection reports >500 plots etc – reports issued within target
- W23.1 Self lay design and terms request <500 plots etc – quotations within target
- W24.1 Self lay design and terms request >500 plots etc – quotations within target
- W26.1 Self lay water for pressure/bacteriological testing – provided within target
- W27.1 Self lay permanent water supply – provided within target
- W30.1 Self lay plot references and costing details – issued within target
- S1.1 Pre-development enquiry – reports issued within target
- S3.1 Sewer requisition design – offers issued within target
- S4.1 Sewer requisition – constructed and commissioned within agreed extension
- S7.1 Adoption legal agreement – draft agreements issued within target
Ofwat (The Water Services Regulation Authority) is a non-ministerial government department. We regulate the water sector in England and Wales.