Appendix 5i:

Understanding Customer Values_ Behavioural Experiment Report

YorkshireWater

PR19 Understanding Customer Values: Work Package 5 – Behavioural Experiment

Prepared for Yorkshire Water

Annex 1 Acknowledgements

Annex 2 AECOM and London Economics would like to thank Professor Nicholas Hanley (University of St. Andrews) for providing peer review of the work undertaken for this work package.

Annex 3 Quality information

Document name	Prepared for	Prepared by	Date	Approved by
Work Package 5	Yorkshire Water	James Suter	01/11/17	Chris White
Work Package 5	Yorkshire Water	James Suter	30/11/17	Chris White

Annex 4

1. Contents

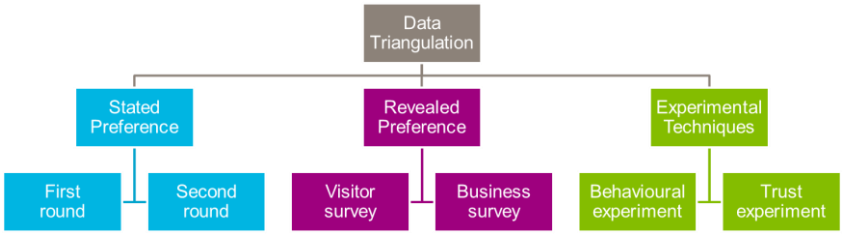
1 Work Package 5 – Behavioural Experiment

2 Context

The aim of this project is to undertake primary research to ascertain the values that Yorkshire Water (YWS) customers place on changes in service measures such as interruptions to supply or drinking water failures. These values will then be used to populate the Decision Making Framework (DMF) in order to inform the investment planning process and support the wider Outcome Delivery Incentives (ODI) work stream.

In light of Ofwat's recommendations for improving the approach to understanding customer's values in PR19, the project includes six work packages (see Figure 1) which draw on a range of data to allow methodological triangulation; whereby data of different types are used to cumulatively refine and validate research outputs.

Figure 1. Overview of the six work packages



This work package seeks to estimate the values YWS' customers place on changes in service measures using a non-incentivised behavioural experiment. One criticism from Ofwat of previous willingness-to-pay (WTP) analysis is that estimates based on stated preference (SP) techniques can be sensitive to the framing of the questions. A behavioural experiment – such as the one implemented in this work package – can test the impacts of such framing effects on results, as well as providing another data source for triangulating values. However, note that, unlike many other experiments, this experiment did not 'incentivise' respondents' behaviour (i.e. respondents' choices did not have real life consequences, such as payments of real money or changes to their actual service levels), as there was no clear way to align the incentives with the desired behaviour of revealing their WTP.

3 Aims

The aim of this work package is to explore YWS customers' preferences over different service levels and the bill impacts using an online behavioural experiment. The purpose of this is to support the triangulation of WTP measures for the DMF and to test the impacts of alternative ways of presenting (or 'framing') the choices presented to respondents. For each of the service areas tested, the amounts by which customers say they want to see bills increase (or decrease) in order to improve (or worsen) service quality are examined using an online instrument – the 'behavioural experiment'. These bill changes are examined for all customers as well as by customer type, based on socio-economic group, age, lifestyle, vulnerability, and income.

The experiment also aims to explore 'framing effects' by testing the impact of three alternative

'treatments' on the amounts that customers say they want to see their bills change by. The alternative treatments are:

- Showing respondents the impact of bill changes on their disposable income.
- Showing respondents comparative information on industry average service levels.
- Changing how the likelihood that unlikely events will occur is presented.

The impact of different costs for adjusting service quality levels (e.g. whether the cost of reducing the number of properties per year affected by low pressure from 15 properties to 1 property would cost £1.00 or £1.30 on every customer's bill) is also tested.

4 Method

This work package provides a further source of WTP values from an exercise that is very different and more interactive and deliberative compared to the standard SP surveys in Work Packages 1 and 2. This enriches the evidence base and provides a further set of values for the triangulation process. It also tests different framing treatments (described further below), complementing the choice experiments carried out in other work packages (our baseline treatment mirrored the choice experiment in Work Package 1, whereas the alternative treatments tested specific adaptations to this baseline approach).

The behavioural experiment took the form of an interactive online tool, which allowed participants to adjust service levels and observe, in real time, the effects that this has on their bill. The operation of the experiment was explained to participants and they were asked to make hypothetical choices regarding their bills and the service levels they would receive from 2020 onwards. They were also provided with details explaining each of the service level attributes and YWS's current performance on those attributes.

Respondents were able to adjust service levels for the same 13 service attributes that were examined in the SP surveys (see Work Packages 1&2). The attributes are categorised into four groups (water quality; supply of water; sewerage services; and environment), with each group of attributes presented on a different screen within the experiment. Respondents used a 'slider' on the screen to adjust the level of service for each attribute. As they moved each slider to increase or decrease the service level, they were shown in real time the impact that this would have on their bill. Respondents also answered questions about themselves before and after the exercise, and were presented with the aggregate impact on their bill of all the changes they had made, at which point they could then to go back and adjust their choices in order to eventually come to the mix of service levels (and corresponding bill) matching their preferences.

Participants were allocated at random to one of four treatment groups. One group received the 'baseline' treatment, while the rest received one of three alternative treatments in which: 1) their remaining disposable income was displayed on screen; 2) industry averages were displayed for comparison for some attributes; and 3) attribute service levels involving low probabilities were presented as frequencies instead of quantities (e.g. instead of *how many properties are affected per year*, respondents were told *every how many hours a property is affected*). Participants were also allocated at random to one of three cost level settings.

The bill increases/decreases that respondents saw as they moved the sliders depended on the cost level setting they were assigned to, and also on their current bill level. The bill impacts were set in line with each respondent's current bill. This meant that the bill increases for a given service improvement were the same across respondents **in percentage terms** (rather than those with lower bills seeing much larger changes in percentage terms), but varied **in absolute terms**.

Data was collected by hosting the online experiment on YouGov's survey platform and drawing the sample from YouGov's panel. A sample of 2,027 respondents living in the YWS area completed the

experiment in late September and early October 2017. Representativeness of YWS's customer base was ensured by setting demographic quotas covering age, gender, sub-region of the YWS area, and social grade so as to match the quotas for the SP surveys conducted in Work Packages 1&2.

5 Results

For all attributes, the results show that a high share of respondents chose the 2020 bill and service level combination, i.e. they chose not to move the sliders. Across the attributes the proportion who did **not** move the slider for that attribute ranged from 35% to 50%. It is not possible to be sure why this is the case. It may be a genuine preference, which could be symptomatic of YWS' current plans being well calibrated to their customers' priorities. It is also possible that it reflects a 'protest vote', status quo bias, respondent fatigue when manipulating the sliders, and/or respondents simply not engaging properly with the choices.

All the results presented in this report are based on 1,736 respondents (86% of the total) – this is all respondents *except those that did not move any sliders at all during the experiment*. Those who did not move any slider for any attribute are not included in the analysis presented here since they are effectively acting like protest bidders – it is not possible to interpret their responses as revealing their demand for the service levels.

An analysis was undertaken to understand which types of respondent were more likely to not move any sliders at all during the exercise and there does not appear to be any link with income. The characteristic that is most strongly related to not moving any sliders, however, is being in the oldest age groups (70 or over). These respondents were less likely to move any sliders. A possible explanation for this is that they had physical or visual difficulty moving the sliders, which would imply that the status quo choice was not necessarily their preference (supporting the argument for omitting these respondents from the analysis).

Baseline results

The SP surveys examined respondents' choices between different levels of service and different **increases or decreases** in their bills and are therefore especially useful for estimating per unit WTP values. The behavioural experiment in this work package prominently presented each respondent with the total bill corresponding to their choices. This means that the results from this work package provide a useful insight into whether respondents would like to increase their total bill above the level implied by service level targets for 2020 in exchange for service level improvements. Table 1 shows that, under the baseline treatment, customers did on average prefer higher bills and higher levels of service relative to the current 2020 targets.

Table 1. Total bill chosen, by starting 2020 bill level

Annex 6 Starting 2020 bill level	Annex 7 £136.50	Annex 8 £405.50	Annex 9 £668.50	Annex 10 932.50	Annex 11 1,195.50	Annex 12 ll
Annex 13 verage (mean)	Annex 14 142	Annex 15 420	Annex 16 691	Annex 17 959	Annex 18 1,230	Annex 19 413
Annex 20 inimum	Annex 21 131	Annex 22 388	Annex 23 638	Annex 24 901	Annex 25 1,155	Annex 26 131
Annex 27 aximum	Annex 28 163	Annex 29 487	Annex 30 804	Annex 31 1,059	Annex 32 1,327	Annex 33 1,327
Annex 34 elow 2020 starting bill (%)	Annex 35 1.0%	Annex 36 0.6%	Annex 37 3.5%	Annex 38 .8%	Annex 39 .3%	Annex 40 1.1%

Annex 41 qual to 2020 starting bill (%)	Annex 42 .2%	Annex 43 .1%	Annex 44 .0%	Annex 45 .0%	Annex 46 .0%	Annex 47 .1%
Annex 48 bove 2020 starting bill (%)	Annex 49 8.8%	Annex 50 9.3%	Annex 51 6.5%	Annex 52 0.2%	Annex 53 1.7%	Annex 54 8.8%

Figure 1 Note: All per unit cost levels, baseline treatment. The different levels of starting bills in 2020 are based on information provided by participants about their current bill.

For each attribute, participants also chose, on average, to increase chosen bills and to improve the service level as shown in Table 2. The average per unit price for each attribute can be used for comparison with the results of the SP analysis in Work Packages 1&2.

Table 2. Preferred change in service level and average per unit price

Annex 55 measure	Service	Annex 56 v. bill increase	A	Annex 57 e change in service level	Averag	Annex 58 e per unit price	Averag
Annex 59 pressure	Poor	Annex 60 £0.11	+	Annex 61 fewer properties affected per year	0.9	Annex 62 per property	£0.12
Annex 63 quality, biological	Water	Annex 64 £1.39	+	Annex 65 more samples out of 10,000	0.3	Annex 66 per sample	£4.74
Annex 67 quality, aesthetic	Water	Annex 68 £3.56	+	Annex 69 fewer customer contacts per year	513	Annex 70 per 1,000 contacts	£6.94
Annex 71 d interruptions	Unplanne	Annex 72 £0.54	+	Annex 73 fewer properties affected per year	2,637	Annex 74 per 1,000 properties	£0.21
Annex 75	Leakage	Annex 76 £1.58	+	Annex 77 million fewer litres lost per day	29	Annex 78 per million litres	£0.05
Annex 79 restrictions	Water	Annex 80 £0.24	+	Annex 81 fewer years out of 100 with a ban	0.6	Annex 82 per year out of 100	£0.4
Annex 83 sewer flooding	Internal	Annex 84 £1.97	+	Annex 85 fewer incidents per year per year	145	Annex 86 per 100 incidents	£1.36
Annex 87 sewer flooding	External	Annex 88 £0.52	+	Annex 89 fewer incidents per year per year	496	Annex 90 per 100 incidents	£0.11
Annex 91	Odour	Annex 92 £0.29	+	Annex 93 fewer complaints per year	683	Annex 94 per 1,000 complaints	£0.43
Annex 95 water	Bathing	Annex 96 £0.46	+	Annex 97 more beaches at 'Good' or 'Excellent'	0.9	Annex 98 per beach	£0.53
Annex 99 water	River	Annex 100 £2.23	+	Annex 101 percent of rivers in Yorkshire improved	2.2	Annex 102 per % of rivers	£1.01
Annex 103 incidents	Pollution	Annex 104 £0.84	+	Annex 105 fewer (Cat. 3) incidents per year	27	Annex 106 per 100 incidents	£3.16
Annex 107 improved / conserved	Land	Annex 108 £0.27	+	Annex 109 ha of land conserved or improved	5452	Annex 110 per 1,000 hectares	£0.05
Annex 111 change	Total bill	Annex 112 £14.02	+	Annex 113	-	Annex 114	-

Figure 2 Note: All starting 2020 bill levels, all per unit cost levels, baseline treatment. The average per

unit price is the average change in the chosen bill level divided by the average change in the service level. Average chosen bill increase is in £/year.

As expected, average chosen bill increases are higher among those paying higher bills, with average total chosen bill increases ranging from £5.91 to £34.86 across the starting bill levels, compared to the average of £14.02.

Average chosen bill increases among those with characteristics suggesting **vulnerability** vary either side of the overall mean, although in general they are lower than the average bill increases chosen by those without any of these characteristics (+£14.66), as shown in Table 3. Average chosen bill increases are lowest for those aged above 75 (+£8.55) and for those with annual incomes below £10,000 (+£9.14).

Table 3. Average chosen bill increases, by vulnerability characteristic

Annex 115 Vulnerability characteristic	Annex 116 ater quality	Annex 117 upply of water	Annex 118 ewerage services	Annex 119 nviron-ment	Annex 120 otal bill change
Annex 121 Aged 75+	Annex 122 3.28	Annex 123 1.45	Annex 124 1.63	Annex 125 2.19	Annex 126 8.55
Annex 127 Worry about affording bill	Annex 128 4.19	Annex 129 2.05	Annex 130 2.57	Annex 131 3.89	Annex 132 12.76
Annex 133 Can't afford bill	Annex 134 5.45	Annex 135 2.15	Annex 136 3.01	Annex 137 4.16	Annex 138 14.53
Annex 139 Receive help to pay bill	Annex 140 6.03	Annex 141 3.26	Annex 142 2.71	Annex 143 2.56	Annex 144 14.56
Annex 145 English not first language	Annex 146 5.85	Annex 147 1.95	Annex 148 2.93	Annex 149 4.25	Annex 150 14.98
Annex 151 Disabled person in household	Annex 152 4.52	Annex 153 2.36	Annex 154 2.67	Annex 155 3.85	Annex 156 13.39
Annex 157 Receive/received benefits	Annex 158 4.43	Annex 159 2.04	Annex 160 2.57	Annex 161 3.12	Annex 162 12.09
Annex 163 Income below £10,000	Annex 164 3.27	Annex 165 1.45	Annex 166 1.91	Annex 167 2.50	Annex 168 9.14
Annex 169 Any reason	Annex 170 4.74	Annex 171 2.15	Annex 172 2.67	Annex 173 3.68	Annex 174 13.19
Annex 175 Financial	Annex 176 4.57	Annex 177 2.09	Annex 178 2.66	Annex 179 3.64	Annex 180 12.88
Annex 181 Physical	Annex 182 4.43	Annex 183 2.22	Annex 184 2.51	Annex 185 3.64	Annex 186 12.78
Annex 187 Not vulnerable	Annex 188 5.03	Annex 189 2.47	Annex 190 2.82	Annex 191 4.46	Annex 192 14.66

Note: All unit cost levels, all starting 2020 bill levels, all treatments.

Average chosen bill increases tend to vary as expected according to other respondent characteristics: chosen bills increase with income, household size, and socio-economic group. For the age category, chosen bills are highest, on average, for the youngest age groups and tend to decline with age, with a bigger decline at age 70 and over. This is an interesting result to compare with results of the SP surveys, since it may demonstrate how designing SP exercises in different ways can affect how different segments of the population respond.

Alternative treatments

At the aggregate level, the **disposable income treatment** had no statistically significant impact

(when an observation is 'statistically significant', this means that it is unlikely to have happened by chance). This suggests that the baseline treatment results are robust to the presentation of disposable income information. However, there is an effect for those who are vulnerable due to disability. The total chosen bill increase for this subgroup was significantly higher (by £4.50) under this treatment. This may be, for example, because displaying their remaining disposable income reassures them that their water bill will not in fact consume a very large share of their disposable income.

Results for the **comparative industry positions treatment** reveal a higher tendency for participants' chosen service levels to cluster around the industry average when this information is shown. It is important to interpret the bill impact results in the context of whether or not this anchoring effect is present (i.e. whether or not the comparative industry information was presented). The general approach taken in the SP surveys in Work Packages 1&2 was to provide comparative industry information and to instruct respondents to take note of this information. However, the dynamics of the choice experiment approach in Work Packages 1&2 mean that it is less likely that this anchoring effect is present, since respondents make choices between options with specified service levels and prices rather than choosing a service level for each attribute individually.

Reframing the service levels of the relevant attributes as frequencies (rather than quantities) had a statistically significant impact on the **service levels chosen** for a number of different attributes. In particular, a worse level of service was typically chosen when the levels were reframed in this way. However, the service level treatment effects are not matched by statistically significant chosen **bill changes** from the reframing of the service level attributes. There was also evidence that presenting the attributes as frequencies instead of quantities made respondents less able to understand the choices. This is an important finding in the context of the wider research programme, since the survey in Work Package 1 used the quantities presentation, which was better understood. Thus, this result from Work Package 5 supports the decision to present the attributes in the way they were in Work Package 1.

Impact of change in cost of services assumption

When the unit costs of changes in service levels were 30% **lower** than the baseline, the average chosen bill increases were **lower**, though usually by less than 30%. This suggests that when prices were lower participants chose better service levels. Similarly, when the unit costs of changes in service levels were 30% **higher** than the baseline, the average chosen bill increases were **higher**, though again usually less than 30% higher, which suggests that participants chose lower service levels when prices were higher. This appears to show that respondents that were responding in an economically rational manner.

Table 4. Average chosen bill increases, by unit cost level

Annex 193 level	Unit cost	Annex 194 (x0.7)	Low	Annex 195 ne	Baseli	Annex 196 (x1.3)	High
Annex 197 quality	Water	Annex 198 0	£4.0	Annex 199	£4.54	Annex 200 9	£6.0
Annex 201 water	Supply of	Annex 202 3	£1.8	Annex 203	£2.31	Annex 204 6	£2.7
Annex 205 services	Sewerage	Annex 206 1	£2.2	Annex 207	£2.69	Annex 208 1	£3.3
Annex 209 ent	Environm	Annex 210 7	£3.2	Annex 211	£3.92	Annex 212 3	£4.9
Annex 213 change	Total bill	Annex 214 32	£11.	Annex 215 6	£13.4	Annex 216 91	£16.

Note: All starting 2020 bill levels, all treatments.

6 Implications

Using the behavioural experiment in this work package it was possible to prominently present each respondent with the revised total bill corresponding to their choices for each individual attribute. This means that the results provide a useful insight into whether respondents would like to increase their total bill above the level implied by service level targets for 2020 in exchange for service level improvements (they would). This information is additional to the information provided from the choice experiments about WTP for marginal changes in service levels. In addition, the WTP results from the choice experiments can be compared with the bill change for unit changes in service levels that are estimated from the results from this work package. This comparison is possible for the average participant as well as by various sub-groups.

A key part of the behavioural experiment was the application of various treatments to test their impact on the results. This responds to some of Ofwat's critiques of previous WTP analysis, which suggest that the results can be sensitive to the framing of the questions, including being influenced by reference points, to other factors such as competing demand on income and also that customers find it hard to assess what they are willing to pay to reduce the probability of a bad, but unlikely, event from occurring.

The results from this work package suggest that when respondents are presented with information about the impact on their disposable income as they are making their choices, on average there is no statistically significant impact on their choices, though there was an impact for people with disabilities. In respect to the presentation of information on industry average performance, the results found that there was a reference point, or anchoring, effect. However, it seems less likely that this effect would materialise in a choice experiment approach to WTP because of the way in which choices are made between bundles of attributes. Some attributes in a bundle may be close to industry averages and others may be further away from industry averages. Finally, the examination of the framing of probabilities provided less clear results, with a framing effect clear in relation to the choice of service levels, but not present in relation to the chosen bill impacts. However, we did find evidence that presenting service levels as quantities as opposed to frequencies resulted in better understanding among consumers.

8 Appendix 1: Methodology

The aim of this work package is to explore the interactions between service quality and the bill impacts of varying service quality using an online behavioural experiment with Yorkshire Water customers. The purpose of this is to support the triangulation of willingness to pay (WTP) measures for Yorkshire Water's Decision Making Framework (DMF) and to test for the presence of three types of framing effects.

- For each of the service areas tested, the amounts by which customers say they want to see bills increase (or decrease) in order to improve (or worsen) service quality are estimated from an online instrument – the 'behavioural experiment'. These bill changes are estimated for all customers as well as by customer type, based on: socio-economic group, age, lifestyle, vulnerability and income.
- The experiment also aims to explore framing effects by testing the impact of three alternative 'treatments' on the amounts by which customers say they want to see their bills change. The three alternative 'treatments' are:
  - showing respondents the impact of bill changes on their disposable income;
  - showing respondents comparative information on industry average service levels; and
  - presenting information on the likelihood of adverse events in different ways.
- The impact of changes in the estimated costs (+/- 30%) of adjusting service quality levels on the results is also tested.

This behavioural experiment takes the form of an interactive online tool, which allows participants to adjust service levels and observe, in real time, the effect that this has on their bills.

The operation of the experiment is explained to participants and they are asked to make hypothetical choices that would affect their bills and the service levels they receive from 2020 onwards. They are also provided with details explaining each of the service level attributes and Yorkshire Water's current performance on those attributes.

Respondents are able to adjust service levels for the same 13 service level attributes that are used in the WTP survey (WP1 and WP2).

Table 1 Service level attributes

Attribute groups	Attributes
Water quality	Poor water pressure
	Drinking water quality
	Taste, smell & colour of drinking water
Supply of water	Unexpected supply interruptions of 3–6 hours
	Leakage
	Water use restrictions
Sewerage services	Sewer flooding inside properties
	Sewer flooding outside properties
	Smell from sewers & treatment works
Environment	Bathing water quality

	River water quality
	Pollution incidents
	Land conserved or improved by Yorkshire Water

The attributes are categorised into four groups (water quality, supply of water; sewerage services; and environment) and each group of attributes is presented on a different screen.

Participants use an on screen slider to adjust service levels for each attribute and as they move the slider to increase or decrease the service level for each attribute, they see in real time the impact that this has on their bill. The starting position for bill impact is always zero and the starting position for service level reflects current levels of service provided by Yorkshire Water.

Respondents also answer other questions about themselves before and after they adjust the sliders and are given an opportunity to see the aggregate impact on their bill of all the changes they have made and then to go back and adjust their choices.

Participants are randomly allocated into four groups. One group receives the baseline treatment and the others one of the three alternative treatments – disposable income, comparative industry position and low probabilities.

Participants are also independently randomly allocated to one of the three levels of cost – baseline, high (+30%) or low (-30%). The baseline cost assumptions are based on Yorkshire Water's estimates of the way costs are likely to vary with the relevant changes in service quality, though there is uncertainty associated with those cost estimates. The high and low costs are intended to test the sensitivity of our results to these different levels of cost.

9. Data collection

Data was collected by hosting the online experiment on YouGov's survey platform and drawing the sample from YouGov's panel[1]. A sample of 2,000 respondents living in the Yorkshire Water area completed the experiment in late September and early October 2017.

This sample is representative of Yorkshire Water's customer base in terms of key demographic characteristics. Representativeness was ensured by setting demographic quotas covering age, gender, sub-region of the Yorkshire Water area and social grade so at to match the quotas for the WTP survey conducted as WP1 and WP2.

Table 2 Target sampling frame

Characteristic	Target
Male	49%
Female	51%
SEG = ABC1	48%
SEG = C2DE	52%
North	16%

[1] YouGov has an active panel of 800,000 potential survey respondents in the UK.

West	45%
South	27%
East	12%
18-34	18%
35-44	17%
45-54	20%
55-64	17%
65+	28%

10.1 Structure of the online tool

The questions that participants face in the online tool are structured as follows.

Introductory and screening questions

Participants are asked for information about whether they have sole or joint responsibility for paying their household water bill, their age, which area of Yorkshire they live in, the occupation of the main income earner in the household, and whether they have a water meter on their property.

Those who do not have sole or joint responsibility for paying their household water bill and those who do not know whether or not they have a water meter are screened out of the process.

Introduction to Yorkshire Water

After random allocation of remaining participants to treatments and price levels (see discussion later), all participants are provided with introductory information about Yorkshire Water and about the water cycle.

Preliminary questions

In order to inform the behavioural experiment, some preliminary questions are asked about current water bills and, for participants allocated to the disposable income treatment only, about disposable income.

Answers to the questions about current water bills determine the 'starting bill' figures presented to participants during the behavioural experiment.

Answers to the questions about disposable income determine the 'starting disposable income' figures presented to participants during the behavioural experiment.

Experiences

Participants are asked questions about their previous experiences in relation to incidents of poor service quality (e.g. discoloured water, sewer flooding, low water pressure etc.); and about whether they have previously visited Yorkshire rivers or coastline for recreational purposes.

Behavioural experiment

The main part of the online tool consists of information presented to participants about the 13 service quality attributes and a series of choices that each participants makes in relation to preferred combinations of water bill and service quality level for each attribute. Further details of the experiment, under each of the four treatments, are provided in subsequent sections of the text.

Review screen

Participants are provided with a summary of the bill impacts related to their previous choices for each of the service level attributes and given the opportunity to go back and change their choices in the behavioural experiment.

Validity questions

Participants are asked whether they felt able to make the choices in the behavioural experiment in a realistic way and if not, why not.

Vulnerability questions

Participants were asked a number of questions in order to assess their vulnerability across a range of dimensions including: ability to afford their water bill, English language ability, disability, receipt of social security benefits and household income.

10.2 Baseline treatment

This the treatment against which the impacts of all other treatments are measured. Hence, under the 'baseline' treatment respondents are not shown the impact of their choices on their disposable income, the comparative industry position of Yorkshire water is not displayed, and service levels that involve low probabilities take a baseline style of presentation.

In order to introduce the choices that participants needed to make, they were presented with an introductory screen explaining the behavioural experiment, followed by a series of screens explaining the service quality indicators for which they would be making choices.[2] These screens were presented by attribute group, such that each screen on which participants made choices for an attribute group was preceded by a series of screens explaining each of the attributes.

The introductory screen, followed by the screen presenting the poor water pressure attribute (from the water quality group of attributes) and the choice screen for water quality are shown in the figures below.

[2] This structure and the screens explaining the service quality indicators were the same across all treatments.

Figure 3 Screenshot showing introduction to choices – baseline treatment

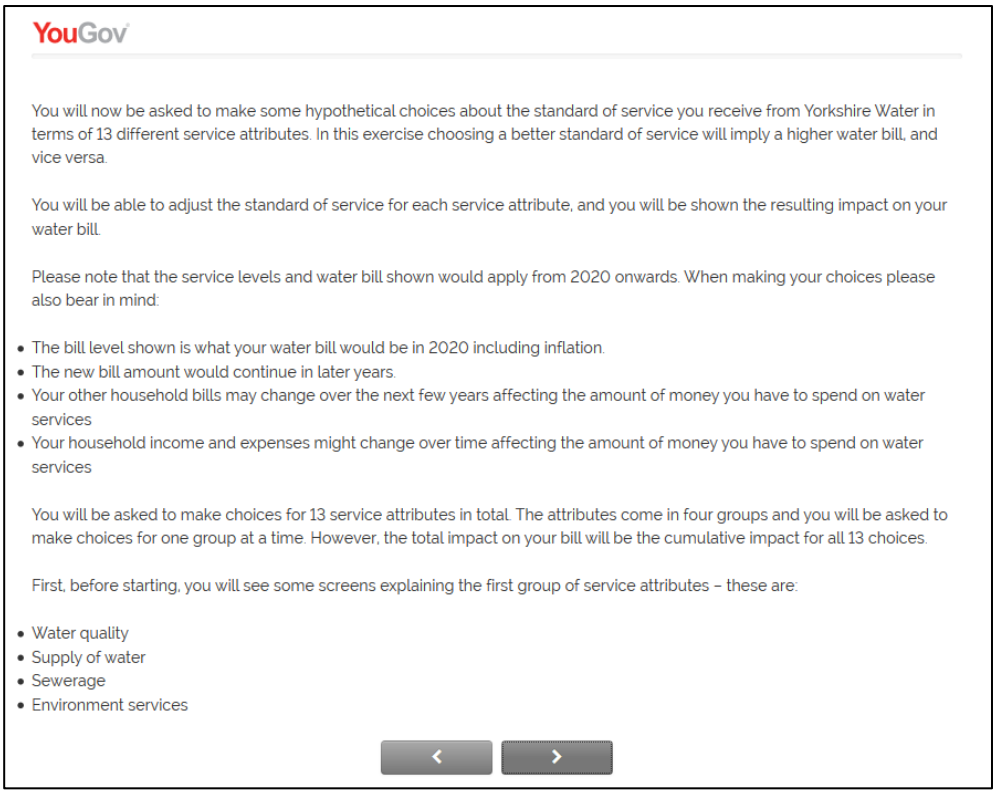


Figure 4 Screenshot showing description of poor water pressure attribute – all treatments

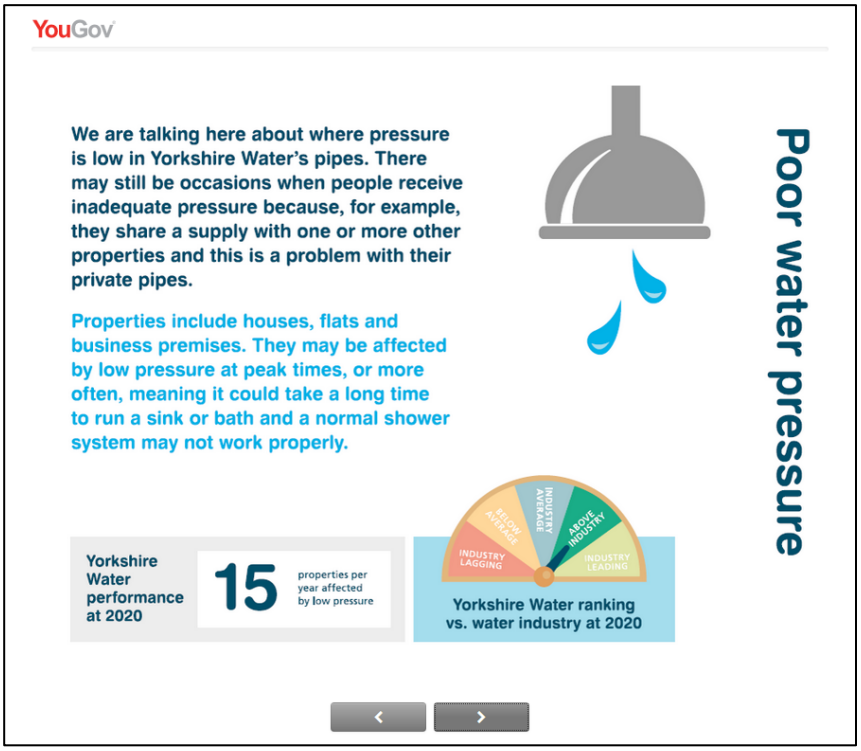
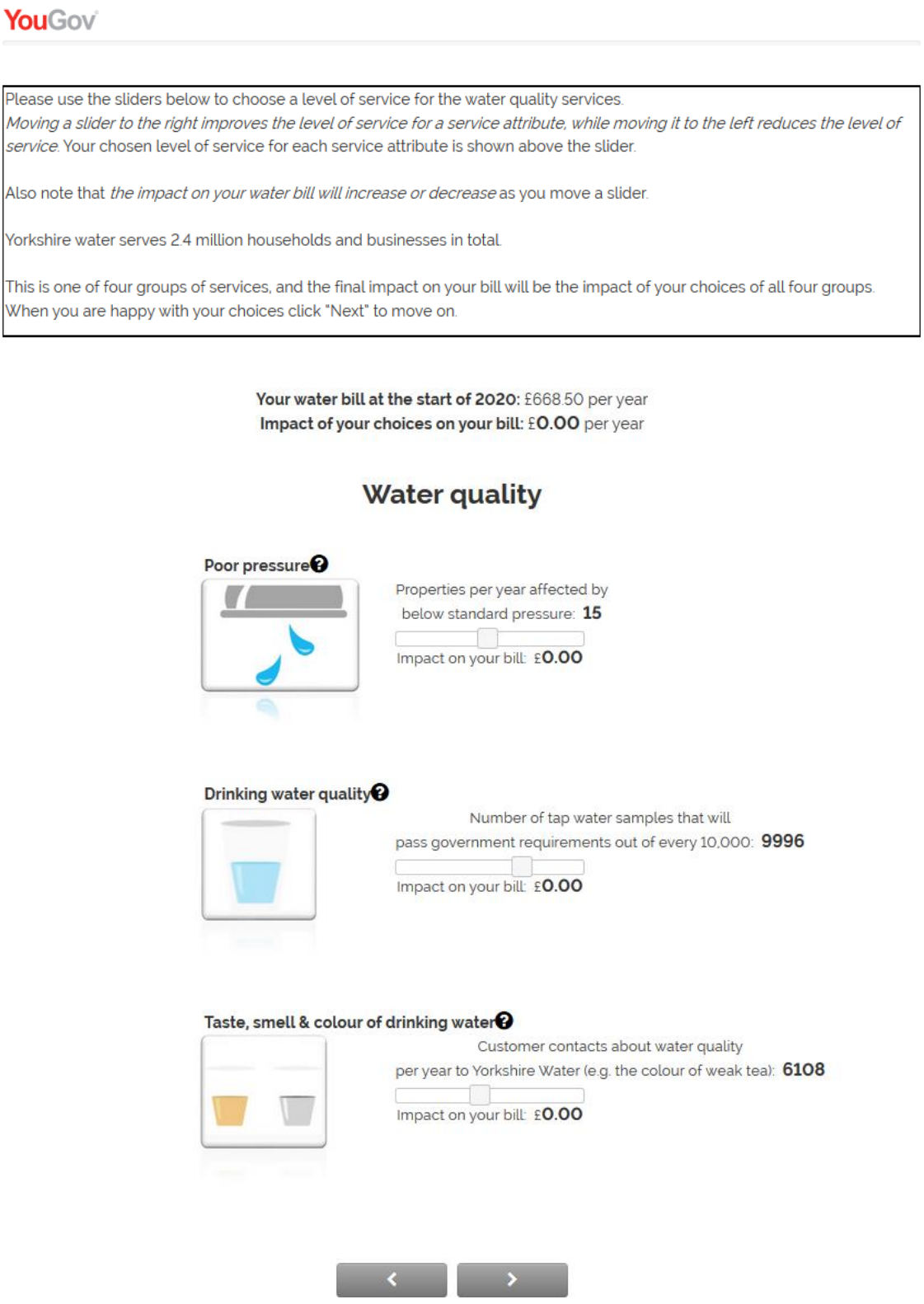
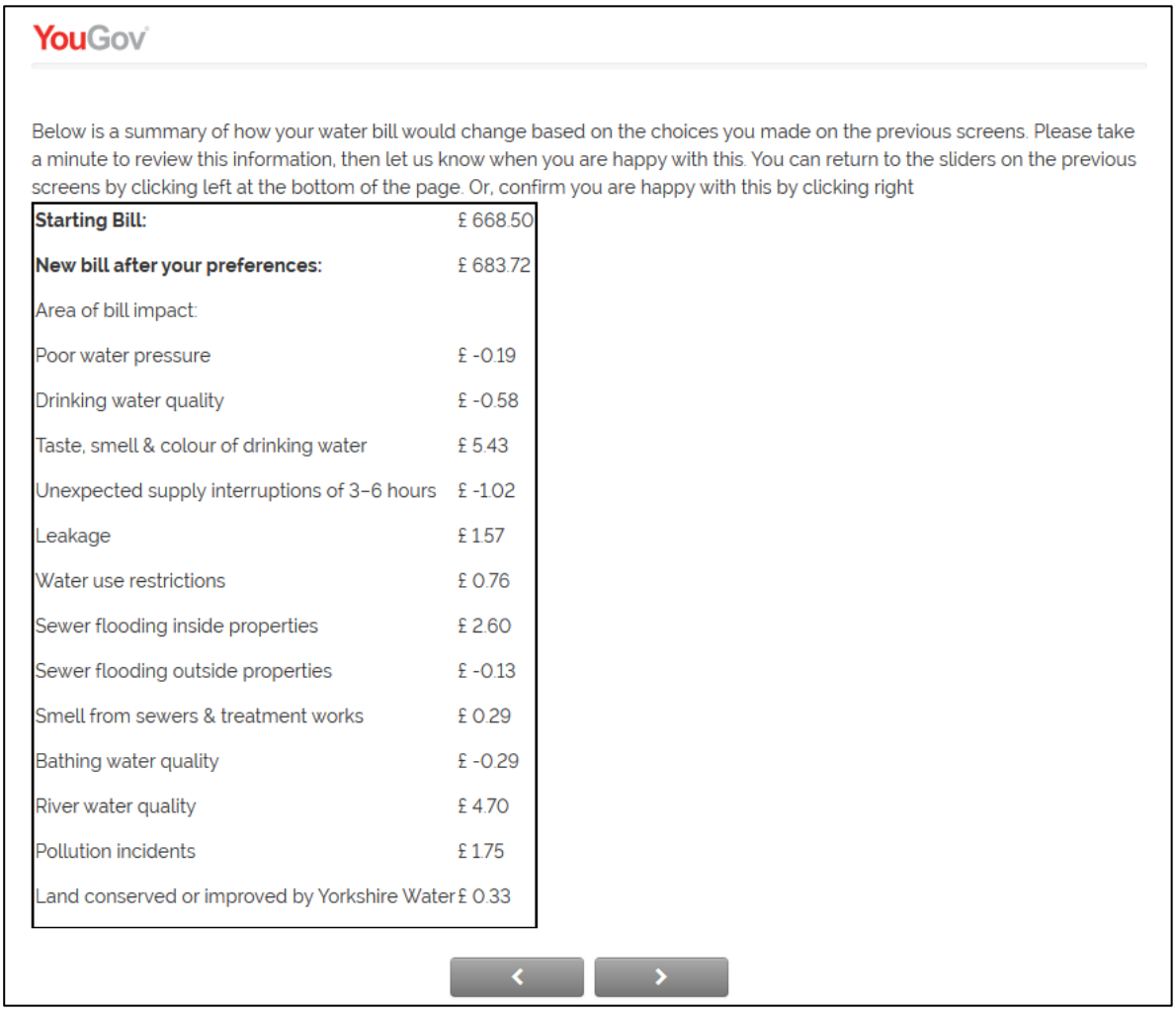


Figure 5 Screenshot showing water quality choices – baseline treatment



After participants have made all the choices in the behavioural experiment (one choice for each of the 13 attributes) they are provided with a summary of the bill impacts related to their previous choices and given the opportunity to go back and change their choices in the behavioural experiment. This review screen is shown below.

Figure 6 Screenshot showing summary of bill impacts – baseline treatment



10.3 Alternative treatment 1: Showing impact on disposable income

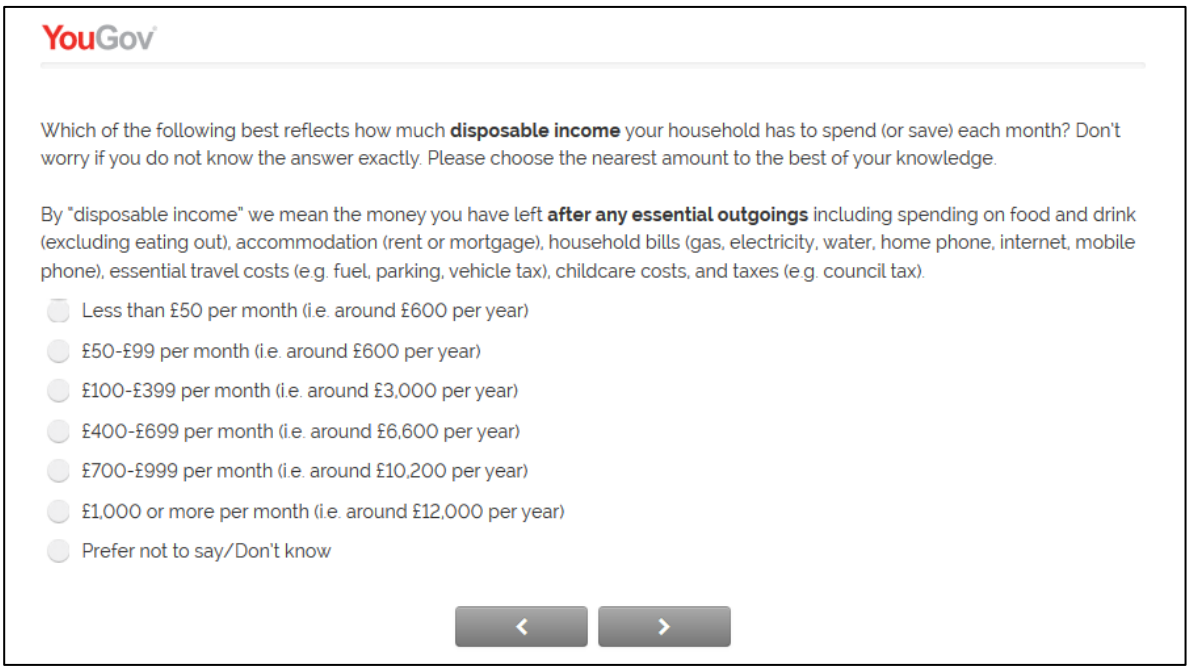
Stated preference (SP) surveys sometimes fail to consider how an individual's income may affect their WTP. Ofwat have previously noted that the values that customers ascribe to particular outcomes in SP surveys may depend on other circumstances beyond the scope of the survey, such as income levels and competing demands on income. That is, valuations from SP surveys may not accurately reflect true valuations if the impacts of respondents' choices on their incomes are not properly articulated in the survey. Thus, our first alternative treatment aims to make the impact of respondents' choices on their remaining discretionary income more salient.

Respondents subject to this treatment are shown their remaining disposable income on the screen alongside their bill, and both of these figures automatically update as the respondent changes their chosen service levels via the sliding bars on the screen.

Whereas there has been substantial debate over the impact of the hypothetical nature of SP surveys on WTP estimates from these surveys[3], this treatment may be expected to reduce any such bias by clearly reminding respondents of their actual budget constraints and helping them to perceive the choices as real economic trade-offs. We anticipate that this treatment is likely to be especially effective for low income respondents in particular who pay more of their income on water services each month.

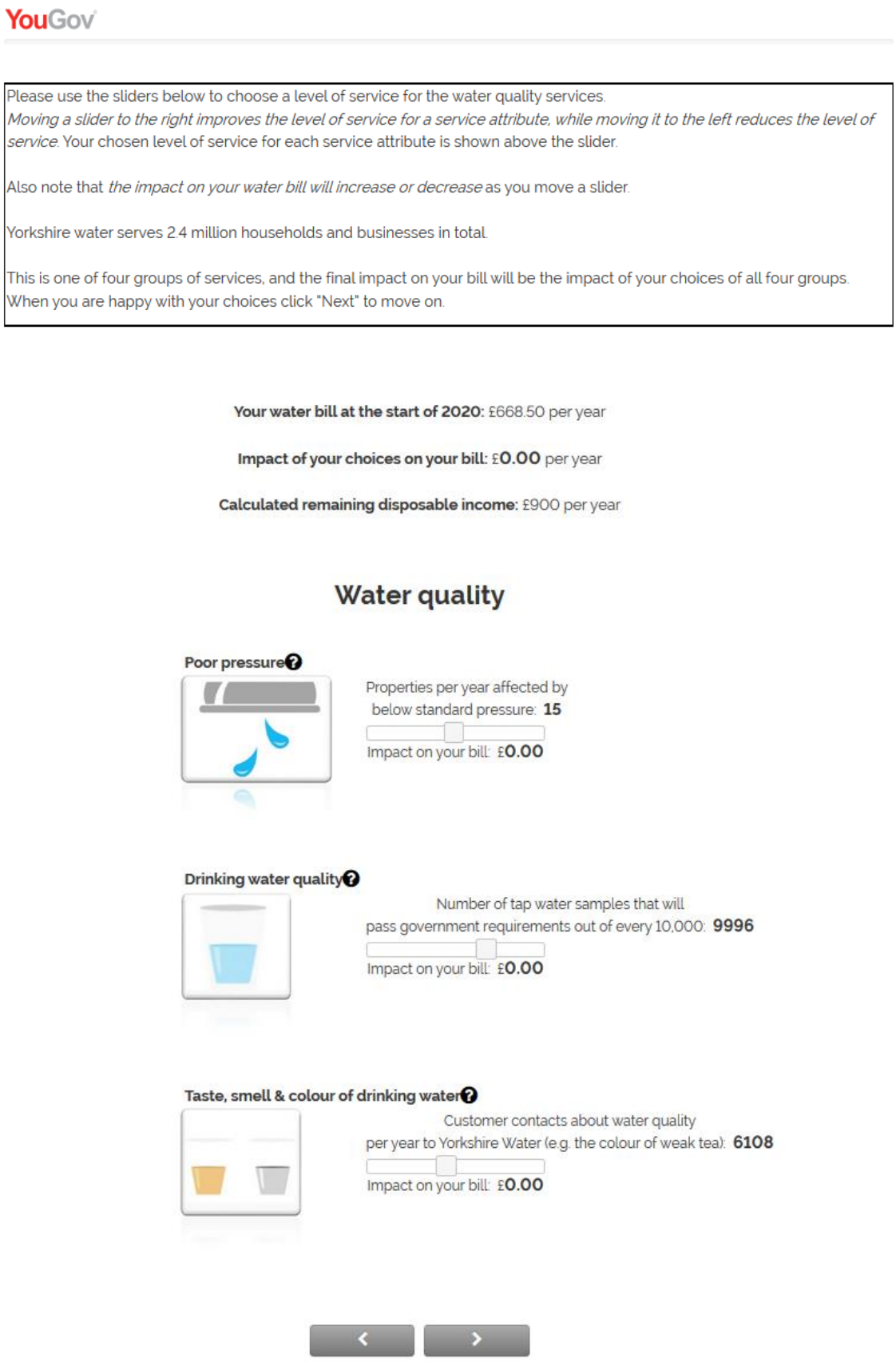
In addition to changes on the screen at which participants made their choice, they were asked a question about their disposable income prior to the screen introducing the behavioural experiment. This is shown below, followed by the screen presenting water quality choices for this treatment.

Figure 7 Screenshot showing disposable income question – disposable income treatment



[3] See Johnston et al (2017), *'Contemporary Guidance for Stated Preference Studies'* (Journal of the Association of Environmental and Resource Economists 2017 4:2, 319-405), who note that there has been substantial recent debate over whether SP methods can provide credible information to inform decision making, with particular attention given to the issue of hypothetical bias

Figure 8 Screenshot showing water quality choices – disposable income treatment



10.4 Alternative treatment 2: Comparative industry information

One criticism from Ofwat of previous WTP analysis is that the estimates can be sensitive to the framing of the questions. In particular, they may be influenced if a reference point is provided in the questionnaire. Extensive research has shown that people tend to rely too heavily on the first piece of information provided to them when making choices.[4] In other words, they compare potential scenarios relative to a baseline, which serves as a mental anchor. As a result, WTP estimates might be reference-dependent.

In order to examine this issue we include a treatment that presents respondents with information on industry average service levels for comparison for particular attributes, in order to investigate how this affects their choices. This information on industry averages was provided on the choice screens for four individual attributes:

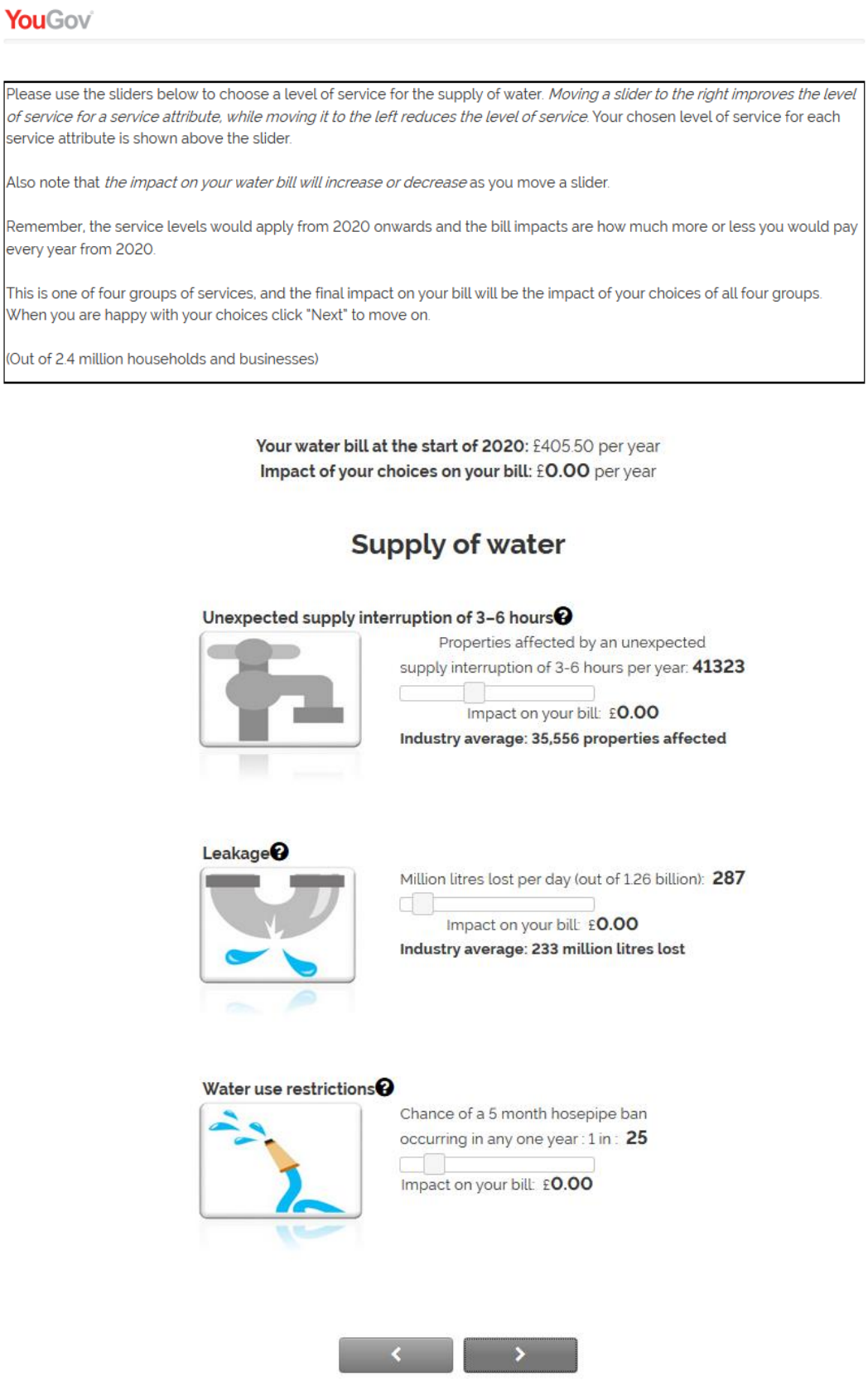
- Supply interruptions
- Leakage
- Internal sewer flooding
- Pollution incidents

This comparative information was not presented for other attributes because it was not readily available for any other attributes.

The choice screen for the supply of water group of attributes under the comparative industry positions treatment is shown below, with the industry averages shown for the supply interruptions and leakage attributes.

[4] For a review of the literature see Furnham and Hua Chu Boo (2011).

Figure 9 Screenshot showing supply of water choices – comparative industry positions treatment



10.5 Alternative treatment 3: Reframing of service attributes involving probabilities

A further issue previously raised by Ofwat in relation to standard SP surveys is that customers find it hard to assess what they are willing to pay to reduce the probability of a bad, but unlikely, event from occurring. This is supported by scientific research in other areas which shows that individuals find it difficult to interpret the low probabilities of bad outcomes (Camerer and Kunreuther, 1989; Magat et al. 1987).

Furthermore, empirical studies have shown that the ways in which probabilities are presented influence decision making, which, in turn, implies that WTP estimates from surveys are dependent on the framing of probabilities. For instance, Stone et al. (1994) find that people are willing to pay considerably more to reduce risk if the likelihood of the event is presented as ratios of very small probabilities rather than if they are presented with two small probabilities.

To investigate this we include a treatment that 'frames' the service levels for particular attributes in alternative ways. The treatment could be applied to those attributes that involve (either implicitly or explicitly) low probabilities of bad events.

A number of potential alternative approaches could be taken for this treatment. These all consist of framing low probabilities in different formats. These are described in the table below (with motivation from the academic literature provided in the footnotes).

It is important to note that not all of these alternative approaches can be applied to very attribute, and that the SP survey in Work Packages 1 and 2 uses a mix of these approaches. This is explained further below.

Table 3 Potential treatments for framing service attribute levels

Framing	Description	Example
Framing 1	Number of incidents or properties affected per year.	"**1,500 incidents** of sewage flooding inside properties **per year**"
Framing 2	Increase the interval over which the number of incidents is counted from 1 year to 5 years (meaning the count is higher).[5]	"**7,500 incidents** of sewage flooding inside properties **every 5 years**"
Framing 3	Present probabilities by dividing by the number of household customers (2.1m).	"**1 in 1,600 chance** of sewage flooding inside any given property over the course of a year"
Framing 4	Present the frequency of an incident, i.e. the average time interval between incidents.[6]	"**One incident** of sewage flooding inside a property **every 6 hours**"

The table below shows which of the alternative framing approaches (from the table above) can be applied to the different service attributes. The table also shows which approaches are used in the SP survey in Work Packages 1 and 2 via the green shading of the cells.

[5] Slovic et al. (1978) find that people exposed to single trips or lifetime (40,000 trips) statistics regarding car accidents and seat belt usage react differently, with the latter condition triggering people to be more willing to favour mandatory protection.

[6] The literature argues that the frequency of bad events conveys the difference in risk more dramatically than the risk probabilities (Weinstein et al., 1996).

Table 4 Which framing options can be applied to which attributes

Attribute	SP survey units and status quo level		Frame 1 Incidents per year	Frame 2 Incidents per 5 years	Frame 3 'X' in 'Y' chance	Frame 4 1 incident every 'X'
Poor Pressure	Properties per year affected by below standard pressure:	15	✓	✓	✓	✓
Drinking Water Quality (Bio)	Proportion of tap water samples that will pass government requirements:	9,996 of 10,000			✓	✓[1]
Drinking Water Quality (Aesthetic)	Customer contacts about water quality per year to Yorkshire Water:	6,108	✓	✓	✓[2]	✓
Unplanned Interruptions	Properties affected by an unexpected supply interruption per year:	41,323	✓	✓	✓	✓
Leakage	Million litres lost per day:	287				
Water Restrictions	Chance of a 5 month hosepipe ban occurring in any one year:	1 in 25			✓	✓
Internal Sewer Flooding	Incidents of sewage flooding inside properties per year:	1,919	✓	✓	✓	✓
External Sewer Flooding	Incidents of sewage flooding outside peoples' homes but on their land, per year:	10,487	✓	✓	✓	✓
Odour	Total complaints per year about unbearable smell from sewerage works:	6,075	✓	✓	✓[3]	✓
Bathing Water Q	Beaches at good or excellent standard:	15 (of 19)				
River Water Q	Percentage of rivers improved:	0%				
Pollution Incidents	Number of minor pollution incidents per year:	211	✓	✓		✓
Land Improved	Hectares of land improved by Yorkshire Water:	0				

Note: Green shading indicates which approaches will be used in the SP survey in Work Packages 1 and 2. [1] "1 sample in 'XXXX' will fail government requirements". [2] "'X' in 'Y' chance any given customer needs to contact Yorkshire Water about their water quality over course of a year". [3] "'X' in 'Y' chance any given customer needs to complain to Yorkshire Water about smell from sewerage works over the course of a year".

The various framings in Table 3 above could be combined in a very large number of ways across the different attributes, and it is not sensible to attempt to examine all of these alternatives. Our approach is to mirror the approach taken in the SP survey in Work Packages 1 and 2 (identified via the green shaded cells in Table 4) for the baseline treatment, and to apply Framing 4 (i.e. present the frequency of an incident) to all relevant attributes as the alternative treatment in this area.

We propose to focus on Framing 4 from Table 3 since for many attributes this presentation is quite different to that which will be used in the SP survey, and because the academic literature argues that the frequency of a bad event conveys the risk more dramatically than the risk probabilities themselves (Weinstein et al., 1996).

The Table below shows the baseline framing for each attribute alongside the framing used in this alternative treatment.

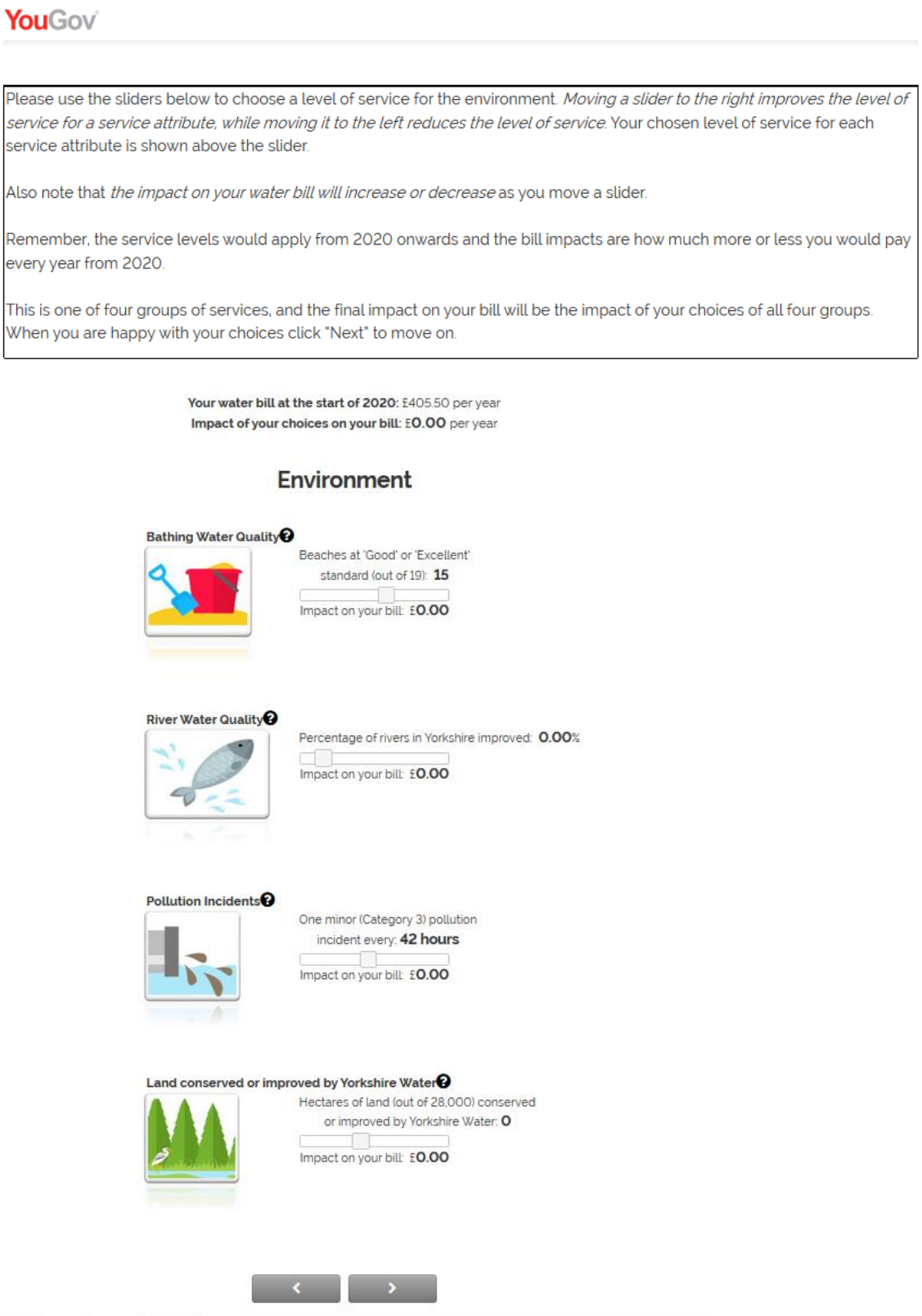
Table 5 Comparison of baseline and alternative treatment 3 framing for each attribute

Attribute	Baseline framing	Alternative framing
Poor water pressure	Properties per year affected by below standard pressure: **XXX**	One property is affected by below standard pressure every **XXX**
Drinking water quality	Share of tap water samples that pass government requirements without health impacts: **XXX**	Tap water samples that fail government requirements with possible health impacts: **XXX**
Taste, smell & colour of drinking water	Customer contacts about water quality (e.g. water the colour of weak tea) per year: **XXX**	One contact about water quality (e.g. water the colour of weak tea) every **XXX**
Unexpected supply interruptions of 3–6 hours	Properties affected by an unexpected supply interruption of 3–6 hours per year: **XXX**	One property affected by unexpected supply interruption of 3–6 hours every **XXX**
Leakage	Million litres lost per day: **XXX**	Million litres lost per day: **XXX**
Water use restrictions	Chance of a 5 month hosepipe ban occurring in any one year (May–Sept): **XXX**	Chance of a 5 month hosepipe ban occurring in any one year (May–Sept): **XXX**
Sewer flooding inside properties	Incidents of sewer flooding of living areas inside properties per year: **XXX**	One incident of sewer flooding of living areas inside a property every **XXX**
Sewer flooding outside properties	Incidents of sewage flooding outside peoples' homes but on their land per year: **XXX**	One incident of sewage flooding outside a person's property but on their land every **XXX**
Smell from sewers & treatment works	Complaints per year about chronic smells from sewers and treatment works: **XXX**	One complaint about chronic smells from sewers and treatment works every **XXX**
Bathing water quality	Beaches at 'Good' or 'Excellent' standard (out of 19): **XXX**	Beaches at 'Good' or 'Excellent' standard (out of 19): **XXX**
River water quality	Percentage of rivers in Yorkshire improved: **XXX**	Percentage of rivers in Yorkshire improved: **XXX**
Pollution incidents	Number of minor (Category 3) pollution incidents per year: **XXX**	One minor (Category 3) pollution incident every: **XXX**
Land conserved or improved by Yorkshire Water	Hectares of land (out of 28,500) conserved or improved by Yorkshire Water: **XXX**	Hectares of land (out of 28,500) conserved or improved by Yorkshire Water: **XXX**

Note: **XXX** is a figure representing the probability in a way consistent with each treatment. The actual figures presented vary as participants move the sliders on the choice screens.

The screenshot on the next page shows how the alternative framing was presented in the choice screen for the environment attributes.

Figure 10 Screenshot showing environment choices – low probabilities framing treatment



10. Appendix 2: Results

11.1 Level of engagement with the experiment

For all attributes we found that a high share of respondents chose the 2020 status quo bill and service level combination, i.e. they chose not to move the sliders – shown as 'Chose 2020 status quo' in Table 6 below. Across the attributes the proportion who did not move the slider for that attribute ranged from 35% to 50%.

Table 6 Proportion of respondents who did not move sliders, by attribute: all respondents

	Chose 2020 status quo	Mean bill increase*
Poor pressure	49%	0.10
Water quality, biological	50%	1.19
Water quality, aesthetic	39%	3.05
Unplanned interruptions	45%	0.47
Leakage	33%	1.36
Water restrictions	47%	0.21
Internal sewer flooding	36%	1.69
External sewer flooding	43%	0.45
Odour	44%	0.25
Bathing water	46%	0.39
River water	35%	1.91
Pollution incidents	41%	0.72
Land improved/conserved	39%	0.23
Total	n/a	12.02

Note: *Mean bill increase under the baseline treatment. Total number of observations: **2,027**.

It is possible that these status quo choices are genuine preferences, which could be symptomatic of Yorkshire Water's current plans being well calibrated to their customers' priorities. Alternatively, it is also possible these responses may reflect 'protest votes' of some kind: status quo bias; respondent fatigue when manipulating the sliders; and/or respondents simply not engaging properly with the choices. The reasons for not moving sliders are likely to vary across respondents, and we cannot be sure of the reasons for individual respondents.

The average number of sliders that were moved, by attribute group are shown in the table below. In total an average of 7.52 sliders out of 13 were moved.

Table 7 Average number of sliders moved (out of 13): all respondents

	Chose 2020 status quo	Mean bill increase*
Water quality attributes	1.62	4.35
Supply of water attributes	1.75	2.03
Sewerage services attributes	1.77	2.39
Environment attributes	2.38	3.26
Total	7.52	12.02

Note: *Mean bill increase under the baseline treatment. Total number of observations: **2,027**.

In order to investigate this further we examined the data for three types of respondent in terms of the way in which they engaged with the experiment:

- Those who chose the 2020 status quo for all attributes, which we believe is likely to be a signal that they simply clicked through the screens without giving the choices a great deal of thought;
- Those who chose the 2020 status quo for all attributes on a particular screen/in a particular group, which similarly may be a signal that they clicked through that particular screen without much thought; and
- Those who answered "No" to the follow-up question "did you feel you were able to make the choices in these exercises in a realistic way?"

The proportions and numbers of respondents in each of these three categories, as well as the mean chosen bill increases for each subgroup under the baseline treatment, are shown for each attribute in the table below.

Table 8 Proportion of respondents who did not move sliders, with selected respondents excluded

	Exclude if chose status quo for all attributes:		Exclude if chose status quo for all attributes in group:		Exclude if said could not answer in a realistic way:	
	Chose 2020 status quo	Mean bill increase*	Chose 2020 status quo	Mean bill increase*	Chose 2020 status quo	Mean bill increase*
Poor pressure	40%	0.11	28%	0.14	49%	0.09
Water quality, biological	42%	1.39	30%	1.70	50%	1.18
Water quality, aesthetic	29%	3.56	14%	4.35	39%	3.10
Unplanned interruptions	36%	0.54	27%	0.62	45%	0.46
Leakage	21%	1.58	11%	1.81	32%	1.38
Water restrictions	38%	0.24	30%	0.28	47%	0.20
Internal sewer flooding	25%	1.97	10%	2.37	35%	1.68
External sewer flooding	33%	0.52	20%	0.63	42%	0.44
Odour	35%	0.29	21%	0.35	43%	0.25
Bathing water	37%	0.46	27%	0.53	45%	0.41
River water	25%	2.23	12%	2.60	33%	2.00
Pollution incidents	31%	0.84	20%	0.98	40%	0.75
Land improved	29%	0.27	17%	0.32	37%	0.26
Total	n/a	14.02	n/a	16.67	n/a	12.19

Note: *Mean bill increase under the baseline treatment. Total number of observations *if those who chose status quo for all attributes are excluded*: **1,736**. Total number of observations *if those who chose status quo for all attributes in group are excluded*: **1,437 to 1,528** (depending on the attribute). Total number of observations *if those who said could not answer in a realistic way are excluded*: **1,519**.

As can be seen from Table 6 and Table 8, the key results (i.e. the mean bill increases) vary a little depending on whether we use the whole sample or alternatively one of the samples where we exclude some respondents. In particular, the effect of dropping some respondents is that the mean chosen bill increases are a little higher (this is expected because bill increases of zero are being dropped).

The average numbers of sliders moved in the three samples with selected respondents excluded are shown in the table below.

Table 9 Average number of sliders moved (out of 13): selected respondents excluded

	Exclude if chose status quo for all attributes:		Exclude if chose status quo for all attributes in group:		Exclude if said could not answer in a realistic way:	
	Chose 2020 status quo	Mean bill increase*	Chose 2020 status quo	Mean bill increase*	Chose 2020 status quo	Mean bill increase*
Water quality	1.89	5.07	2.28	6.19	1.62	4.37
Supply of water	2.05	2.37	2.33	2.70	1.77	2.04
Sewerage services	2.07	2.79	2.49	3.35	1.81	2.37
Environment	2.78	3.80	3.24	4.43	2.45	3.41
Total	8.78	14.02	10.34	16.67	7.65	12.19

Note: *Mean bill increase under the baseline treatment. Total number of observations *if those who chose status quo for all attributes are excluded*: **1,736**. Total number of observations *if those who chose status quo for all attributes in group are excluded*: **1,437 to 1,528** (depending on the attribute). Total number of observations *if those who said could not answer in a realistic way are excluded*: **1,519**.

We also examined the correlations between whether or not respondents chose the 2020 status quo for each attribute. These are shown in Table 10. In general, respondents who chose the 2020 status quo for one attribute were more likely to choose the 2020 status quo for another attribute as well.

When we present our main results we have elected to focus on the sample of 1,736 respondents that includes all respondents except those that chose the 2020 status quo for all attributes. The respondents that did not move any sliders are more clearly acting like 'protest bidders' than those who did not move sliders for a particular group of attributes. We cannot with much certainty interpret their responses as revealing their demand for the service levels. We preferred to exclude this group, rather than the group who reported in the post experiment question that they could not answer in a realistic way, because we prefer to exclude on the basis of actual behaviour in the experiment, rather than a more subjective view post experiment.

Table 10 Correlations between status quo choices for each of the 13 individual attributes

	Poor pressure	Water quality, biological	Water quality, aesthetic	Unplanned interruptions	Leakage	Water restrictions	Internal sewer flooding	External sewer flooding	Odour	Bathing water	River water	Pollution incidents	Land improved/ conserved
Poor pressure	1.00												
Water quality, biological	0.47	1.00											
Water quality, aesthetic	0.50	0.60	1.00										
Unplanned interruptions	0.42	0.41	0.48	1.00									
Leakage	0.35	0.41	0.50	0.53	1.00								
Water restrictions	0.39	0.40	0.41	0.53	0.53	1.00							
Internal sewer flooding	0.40	0.46	0.51	0.52	0.56	0.47	1.00						
External sewer flooding	0.40	0.42	0.48	0.51	0.51	0.47	0.70	1.00					
Odour	0.38	0.40	0.44	0.51	0.52	0.48	0.60	0.64	1.00				
Bathing water	0.33	0.37	0.39	0.41	0.45	0.38	0.48	0.47	0.46	1.00			
River water	0.30	0.35	0.42	0.42	0.52	0.39	0.51	0.51	0.48	0.60	1.00		
Pollution incidents	0.34	0.38	0.46	0.49	0.54	0.43	0.55	0.57	0.54	0.60	0.65	1.00	
Land improved/conserved	0.30	0.38	0.43	0.41	0.51	0.41	0.52	0.51	0.51	0.59	0.69	0.67	1.00

11.2 Econometric analysis of respondents engagement with the experiment

We undertook econometric analysis to understand which types of respondent were more likely to not move any sliders at all during the exercise. This analysis examines the following explanatory variables together:

- Age
- Household size
- Income
- Socio-economic group

Furthermore, in a separate analysis the following (binary) vulnerability indicators are examined:

- Aged 75+
- Worry about affording bill
- Can't afford bill
- Receive help to pay bill
- English not first language
- Disabled person in household
- Receive/received benefits
- Income below £10,000

The variables representing age and income were each coded in two alternative ways:[7]

- A continuous variables:
  - For age, the variable is in years ranging from 19 to 89 years.
  - For income, we take the mid-points of the bands offered in the survey question.[8]
- As sets of dummies:
  - For age, seven dummies signifying whether the respondent is aged 18-29, 30-39, 40-49, 50-59, 60-69, 70-79, or 80-89 years
  - For income, six dummies signifying whether the respondent has a household income of under £10,000, £10,000-£19,999, £20,000-£29,999, £30,000-£39,999, £40,000-£49,999, or £50,000 or more.

Household size is coded as a continuous variable (taking values 1 to 8, where 8 means '8 or more') and socio-economic group is coded as a set of dummies signifying whether the respondent is in group A, B, C1, C2, D or E.

[7] Coding these as continuous variables provides results that are more parsimonious and straight-forward to interpret. On the other hand, the benefit of coding the variables as sets of dummies is that this allows the relationships between the explanatory characteristics and the dependent variable to be non-linear. For example, the estimated difference between those earning £10,000-£19,999 and those earning £20,000-£29,999 does not have to be the same as the estimated difference between those earning £20,000-£29,999 and those earning £30,000-£39,999. A separate regression coefficient is estimated for each group represented by a dummy, relative to a specified base group.

[8] Those who indicated an income of '£50,000 or more' assigned an income of £55,000. Thus the variable takes the values £5,000, £15,000, £25,000, £35,000, £45,000, and £55,000.

We follow the general to specific approach to specifying our econometric model. That is, we include all variables in the first estimation of the model, then remove any that are not statistically significant and re-run the model, and so on, until we reach a model that only includes statistically significant variables. This approach is followed using the continuous variables for age and income and, separately, using the dummies for these characteristics.

Since the dependent variable is binary a **logit model** is estimated in this analysis. The dependent variable represents whether or not the respondent moved any sliders. Specifically, it is equal to:

- 1 if the represent did not move any sliders; and
- 0 if the represent moved at least one slider

General to specific model specification – age, household size, income and socio-economic group

When all explanatory variables are included in the 'general' model only the age variables are statistically significant – see Table 11. The first variable to be dropped during the general to specific process is income due to the low significance of the results for this characteristic, followed by household size. Once income and household size have been removed from the model, the result for the highest socio-economic group (group A) becomes statistically significant and hence all socio-economic group dummies are retained in the final model (the results of which are below).

Table 11 Regression analysis of 'moved any sliders' on age, household size, income, and socio-economic group – General model results

Model 1: Age and income as continuous variables			Model 2: Age and income as dummies		
Explanatory variable	**Coeff.**	**p-value**	**Explanatory variable**	**Coeff.**	**p-value**
Age (continuous)	-0.02***	0.000	Age: 30-39	-1.00*	0.069
			Age: 40-49	-1.27**	0.019
			Age: 50-59	-1.26**	0.019
			Age: 60-69	-1.63***	0.002
			Age: 70-79	-1.53***	0.005
			Age: 80-89	-1.46**	0.041
Household size	0.06	0.343	Household size	0.07	0.279
Income (continuous)	0.00	0.692	Income: £10,000-£19,999	-0.25	0.344
			Income: £20,000-£29,999	-0.04	0.879
			Income: £30,000-£39,999	0.13	0.671
			Income: £40,000-£49,999	0.02	0.963
			Income: £50,000 or more	-0.11	0.733
SE group A	0.32	0.252	SE group A	0.29	0.303
SE group D	0.45	0.075	SE group D	0.40	0.117
SE group C1	0.08	0.726	SE group C1	0.05	0.829
SE group C2	0.04	0.878	SE group C2	0.01	0.957
SE group D	0.15	0.610	SE group D	0.15	0.616
Constant	2.53***	0.000	Constant	2.90***	0.000

Note: All per unit cost levels, all starting 2020 bill levels, all treatments. The continuous income variable is in 1000s of £s. Where dummies are included, the base groups are 18-29 years, income under £10,000, and socio-economic group (E). ***/**/* signifies statistical significance at 99%/95%/90% level.

General to specific model specification – vulnerability indicators

In the general model with vulnerability indicators as regressors, when all explanatory variables are included only being over 75 and being worried about affording their bill are statistically significant – see Table 12. All other variables in this general model are gradually eliminated in turn, due to non-statistical significance, by the general to specific process (hence only being over 75 and being worried about affording their bill remain in the final model – see below).

Table 12 Regression analysis of 'moved any sliders' on vulnerability indicators – General model results

Vulnerability indicators	Coefficient	p-value
Aged 75+	-0.42*	0.080
Worry about affording bill	0.52***	0.008
Can't afford bill	-0.09	0.752
Receive help to pay bill	-0.20	0.726
English not first language	-0.51	0.184
Disabled person in household	-0.23	0.145
Receive/received benefits	0.16	0.363
Income below £10,000	-0.25	0.299
Constant	1.77	0.000

Note: All per unit cost levels, all starting 2020 bill levels, all treatments. ***/**/* signifies statistical significance at 99%/95%/90% level.

Final regression results

The final results for the model with age and socio-economic group as the explanatory variables are presented in Table 13. The final results for the model with the vulnerability indicators being over 75 and being worried about affording their bill as the regressors are presented in Table 14.

These results show that there is a slight between being in the very highest socio-economic group and being more likely to have moved at least one slider. The results also show that those who said they worry about their bill were in fact more likely to move at least one slider.

The characteristic that is most strongly related to not moving any sliders, however, is being in the oldest age groups (70 or over). These respondents were less likely to move any sliders. A possible explanation for this is that they had physical or visual difficulty moving the sliders, which would imply that the status quo choice was not necessarily their preference (supporting the argument for omitting these respondents from the analysis).

Table 13 Regression analysis of 'moved any sliders' on age and socio-economic group – Final model results

	Logit regression results		Marginal effects	
	Coefficient	p-value	Coefficient	p-value
Age coded as a continuous variable:				
Age	-0.02***	0.000	-0.002***	0.000
SE group A	0.44*	0.061	0.048**	0.035
SE group D	0.33	0.111	0.038*	0.090
SE group C1	0.08	0.697	0.009	0.693
SE group C2	0.11	0.655	0.012	0.645

SE group D	0.20	0.487	0.022	0.460
Constant	2.75	0.000	n/a	n/a
Age coded as dummies:				
Age: 30-39	-0.90*	0.070	-0.13	0.130
Age: 40-49	-1.18**	0.016	-0.18**	0.049
Age: 50-59	-1.21**	0.012	-0.19**	0.037
Age: 60-69	-1.53***	0.001	-0.23***	0.006
Age: 70-79	-1.56***	0.001	-0.26***	0.009
Age: 80-89	-1.44**	0.032	-0.26*	0.098
SE group A	0.43*	0.067	0.05**	0.040
SE group D	0.33	0.115	0.04	0.094
SE group C1	0.07	0.726	0.01	0.723
SE group C2	0.11	0.637	0.01	0.627
SE group D	0.20	0.486	0.02	0.458
Constant	2.88***	0.000	n/a	n/a

Note: All per unit cost levels, all starting 2020 bill levels, all treatments. Where dummies are included, the base groups are 18-29 years, income under £10,000, and socio-economic group (E). ***/**/* signifies statistical significance at 99%/95%/90% level.

Table 14 Regression analysis of 'moved any sliders' on vulnerability indicators – Final model results

	Logit regression results		Marginal effects	
Vulnerability indicators	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Aged 75+	-0.44*	0.063	-0.06*	0.099
Worry about affording bill	0.45***	0.007	0.05***	0.003
Constant	1.72***	0.000	n/a	n/a

Note: All per unit cost levels, all starting 2020 bill levels, all treatments. ***/**/* signifies statistical significance at 99%/95%/90% level.

11.3 Baseline results

The section presents the baseline results. These are based on the data from the baseline treatment (since this was equivalent to the SP survey in WP1&2), and exclude respondents who did not change any service levels away from the 2020 status quo (i.e. who did not move any sliders).

The SP survey (WP1&2) examined respondents' choices between different levels of service and different *increases* in their bills and is therefore is especially useful for estimating per unit WTP values. However, it did not present respondents with the total bill levels associated with their choices. The WP5 behavioural experiment, on the other hand, presented each respondent with the total bill corresponding to their choices and clearly showed respondent how their bill would change depending on the their service level choices. Thus WP5 can provide a useful insight into whether respondents would like to increase their total bill above the level implied by service level targets for 2020 in exchange for service level improvements.

Table 15 shows that, under the baseline treatment, customers did on average prefer higher bills and higher levels of service relative to the current 2020 targets.

Table 15 Total bill chosen, by starting 2020 bill level

Starting 2020 bill level:	£136.50	£405.50	£668.50	£932.50	£1,195.5	All

Average (mean)	£142	£420	£691	£959	£1,230	£413
Minimum	£131	£388	£638	£901	£1,155	£131
Maximum	£163	£487	£804	£1,059	£1,327	£1,327
Below 2020 starting bill (%)	11.0%	10.6%	13.5%	9.8%	8.3%	11.1%
Equal to 2020 starting bill (%)	0.2%	0.1%	0.0%	0.0%	0.0%	0.1%
Above 2020 starting bill (%)	88.8%	89.3%	86.5%	90.2%	91.7%	88.8%

Note: All per unit cost levels, baseline treatment. The different levels of starting bills in 2020 are based on information provided by participants about their current bill.

For each attribute, participants also chose on average to increase chosen bills and to improve the service level. Table 16 below shows the average chosen bill increase per attribute, the average chosen service level, the average change in service level relative to the current 2020 targets, and the average per unit price implied by respondents' choices which can be used for direct comparison with the results of the WTP analysis in WP1 and WP2.

Table 16 Total bill chosen, by starting 2020 bill level

	Average chosen bill increase	Average chosen service level	Average change in service level	Average per unit price
Poor pressure	£0.11	14.1 properties affected per year	0.9 fewer properties affected per year	£0.12 per property
Water quality, biological	£1.39	9996 samples out of 10,000	0.3 more samples out of 10,000	£4.74 per sample
Water quality, aesthetic	£3.56	5595 customer contacts per year	513 fewer customer contacts per year	£6.94 per 1,000 contacts
Unplanned interruptions	£0.54	38686 properties affected per year	2637 fewer properties affected per year	£0.21 per 1,000 properties
Leakage	£1.58	257.8 million litres lost per day	29 million fewer litres lost per day	£0.05 per million litres
Water restrictions	£0.24	3.4 years out of 100 with a ban	0.6 fewer years out of 100 with a ban	£0.4 per year out of 100
Internal sewer flooding	£1.97	1774 incidents per year per year	145 fewer incidents per year per year	£1.36 per 100 incidents
External sewer flooding	£0.52	9991 incidents per year per year	496 fewer incidents per year per year	£0.11 per 100 incidents
Odour	£0.29	5392 complaints per year	683 fewer complaints per year	£0.43 per 1,000 complaints
Bathing water	£0.46	15.9 beaches at 'Good' or 'Excellent'	0.9 more beaches at 'Good' or 'Excellent'	£0.53 per beach
River water	£2.23	2.2 percent of rivers in Yorkshire improved	2.2 percent of rivers in Yorkshire improved	£1.01 per percentage of rivers
Pollution incidents	£0.84	184 pollution incidents (Cat. 3) per year	27 fewer (Cat. 3) pollution incidents per year	£3.16 per 100 incidents
Land improved/conserved	£0.27	5452 ha of land conserved or improved	5452 ha of land conserved or improved	£0.05 per 1,000 hectares
Water quality total	£5.07	n/a	n/a	n/a
Supply of water total	£2.37	n/a	n/a	n/a
Sewerage services total	£2.79	n/a	n/a	n/a
Environment total	£3.80	n/a	n/a	n/a
Total bill change	£14.02	n/a	n/a	n/a

Note: All starting 2020 bill levels, all per unit cost levels, baseline treatment.

11.4 Results by starting 2020 bill level and by unit cost level

This section presents information on the bill increases chosen by respondents *by starting 2020 bill level and by unit cost level*. As explained in the methodology section, respondents by starting 2020 bills varied over five levels, and the per unit costs associated with changes in service levels varied over three levels in the experiment (baseline, low (-30%) and high (+30%)). In this analysis, we average across the treatment variants in order to ensure sufficient sample sizes per subgroup.

As expected, average chosen bill increases are higher (in absolute terms) among those paying higher bills, with average chosen total bill increases for each starting 2020 bill level ranging from £5.91 to £34.86, compared to the overall average of £14.02. See the table below.

Table 17 Average chosen bill increases, by 2020 starting bill level

	£136.50	£405.50	£668.50	£932.50	£1,195.50
Poor pressure	£0.04	£0.12	£0.09	£0.21	£0.15
Water quality, biological	£0.62	£1.53	£2.31	£3.69	£3.82
Water quality, aesthetic	£1.36	£3.29	£5.89	£6.35	£6.88
Unplanned interruptions	£0.24	£0.54	£0.66	£0.61	£1.14
Leakage	£0.66	£1.72	£2.32	£2.43	£4.89
Water restrictions	£0.10	£0.31	£0.42	£0.32	£0.94
Internal sewer flooding	£0.82	£2.12	£2.97	£3.92	£5.78
External sewer flooding	£0.19	£0.55	£0.77	£0.83	£1.19
Odour	£0.12	£0.30	£0.45	£0.32	£0.63
Bathing water	£0.18	£0.48	£0.76	£1.02	£1.37
River water	£1.13	£2.52	£3.85	£4.47	£6.32
Pollution incidents	£0.36	£0.88	£1.41	£1.63	£2.40
Land improved/conserved	£0.12	£0.30	£0.44	£0.36	£0.72
Water quality total	£2.01	£4.94	£8.29	£10.25	£10.84
Supply of water total	£1.01	£2.57	£3.40	£3.35	£6.98
Sewerage services total	£1.12	£2.97	£4.19	£5.07	£7.60
Environment total	£1.80	£4.18	£6.45	£7.48	£10.81
Total	£5.91	£14.65	£22.11	£26.15	£34.86

Note: All per unit cost levels, all treatments.

When the unit costs of changes in service levels were 30% *lower* than the baseline, the average total bill increases chosen by participants were lower, though usually by less than 30%. This suggests that when prices were lower, participants chose better service levels (i.e. there was some responsiveness to price).

Similarly, when the unit costs of changes in service levels were 30% *higher* than the baseline, the average total bill increases chosen were *higher*, though again usually less than 30% higher, suggesting that respondents chose lower service levels when prices were higher.

Table 18 Average chosen bill increases, by unit cost level

Unit cost level:	Low (x0.7)	Baseline	High (x1.3)
Poor pressure	£0.07	£0.11	£0.10
Water quality, biological	£1.18	£1.48	£1.84

Water quality, aesthetic	£2.75	£2.95	£4.15
Unplanned interruptions	£0.40	£0.45	£0.57
Leakage	£1.23	£1.59	£1.86
Water restrictions	£0.21	£0.27	£0.34
Internal sewer flooding	£1.61	£1.97	£2.36
External sewer flooding	£0.40	£0.46	£0.61
Odour	£0.20	£0.27	£0.34
Bathing water	£0.39	£0.45	£0.58
River water	£1.97	£2.38	£2.98
Pollution incidents	£0.70	£0.84	£1.03
Land improved/conserved	£0.22	£0.26	£0.34
Water quality total	£4.00	£4.54	£6.09
Supply of water total	£1.83	£2.31	£2.76
Sewerage services total	£2.21	£2.69	£3.31
Environment total	£3.27	£3.92	£4.93
Total	£11.32	£13.46	£16.91

Note: All starting 2020 bill levels, all treatments.

11.5 Results by socio-demographic group

This section presents results by socio-demographic group, for the following socio-demographic characteristics:

- Age
- Household size
- Income
- Socio-economic group (A, B, C1, C2, D, E)

Average chosen bill increases tend to vary by participants' characteristics in the directions that would be expected. As can be seen from Table 19 below, chosen bill increases rise with income, household size and socio-economic group.

Chosen bills tend to decline as age increases, with a bigger decline at age 70 and over. This is an interesting result to compare with results of the SP survey (WP1 and WP2), since it may demonstrate how designing stated preference exercises in different ways can affect how different segments of the population respond to them.

The younger age groups, especially the youngest group (18-29 years), placed more weight on the water quality attributes and less weight on environment attributes: 42% of the average total bill increase chosen by the youngest group is accounted for by the water quality attributes (i.e. £7.32 out of £17.27), compared to just 21% for the environment attributes; whereas, for the other age groups the water quality attributes and the environment attributes each account for more similar proportions of the total bill increase.

The detailed results for the individual attributes and age groups in Table 20 show that the higher average total bill increase chosen by the 18-29 group is in fact driven in large part by this group choosing higher bill increases for the aesthetic water quality attribute.

Table 19 Average chosen bill increases – Attribute group totals, by socio-demographic group

	Water quality total	Supply of water total	Sewerage services total	Environment total	Total
Respondent age:					
18-29 years	£7.32	£3.00	£3.36	£3.67	£17.27
30-39 years	£6.25	£2.67	£3.48	£4.65	£17.12
40-49 years	£5.50	£2.38	£3.05	£4.85	£15.57
50-59 years	£4.78	£2.38	£2.88	£4.03	£13.86
60-69 years	£4.53	£2.40	£2.53	£4.26	£13.69
70-79 years	£3.05	£1.46	£1.83	£2.74	£9.05
80-89 years	£2.23	£1.40	£1.55	£1.72	£6.90
Household size:					
1 person	£2.93	£1.35	£1.76	£2.83	£8.77
2 people	£4.81	£2.51	£2.63	£4.18	£14.07
3 people	£5.60	£2.21	£3.23	£4.13	£15.16
4 people	£7.15	£3.26	£4.09	£5.16	£19.29
5 + people	£7.04	£2.58	£3.83	£5.45	£19.08
Income:					
Under £10,000	£3.27	£1.45	£1.91	£2.50	£9.14
£10,000-£19,999	£3.72	£1.84	£2.32	£3.39	£11.11
£20,000-£29,999	£4.88	£2.10	£2.54	£3.65	£13.17
£30,000-£39,999	£5.11	£2.49	£2.91	£4.79	£15.18
£40,000-£49,999	£5.01	£2.29	£2.58	£3.78	£13.66
£50,000 or more	£7.85	£3.58	£4.29	£6.33	£21.90
Socio-economic group:					
A	£5.39	£2.83	£3.35	£4.44	£15.97
B	£5.09	£2.37	£2.73	£4.34	£14.39
C1	£5.17	£2.05	£2.72	£4.11	£14.05
C2	£4.87	£2.34	£2.78	£3.95	£13.85
D	£4.97	£2.06	£2.51	£3.37	£12.50
E	£3.14	£2.10	£2.11	£3.32	£10.66

Note: All starting 2020 bill levels, all per unit cost levels, all treatments.

Table 20 Average chosen bill increases, by age group

	Age of respondent:						
	18-29	30-39	40-49	50-59	60-69	70-79	80-89
Poor pressure	£0.20	£0.08	£0.08	£0.12	£0.09	£0.07	-£0.03
Water quality, biological	£2.16	£1.94	£1.68	£1.54	£1.38	£0.87	£1.04
Water quality, aesthetic	£4.96	£4.23	£3.75	£3.12	£3.06	£2.11	£1.22
Unplanned interruptions	£0.93	£0.64	£0.51	£0.49	£0.43	£0.19	£0.02
Leakage	£1.62	£1.72	£1.62	£1.57	£1.71	£1.09	£1.26
Water restrictions	£0.45	£0.30	£0.25	£0.32	£0.26	£0.18	£0.11
Internal sewer flooding	£2.35	£2.53	£2.28	£2.01	£1.79	£1.38	£1.44
External sewer flooding	£0.58	£0.58	£0.51	£0.59	£0.49	£0.29	£0.00
Odour	£0.42	£0.38	£0.26	£0.28	£0.26	£0.15	£0.11

Bathing water	£0.34	£0.49	£0.61	£0.55	£0.48	£0.29	£0.24
River water	£2.11	£2.77	£2.85	£2.33	£2.66	£1.77	£1.15
Pollution incidents	£0.87	£1.06	£1.09	£0.88	£0.84	£0.51	£0.22
Land improved/conserved	£0.35	£0.32	£0.30	£0.27	£0.29	£0.17	£0.10
Water quality total	£7.32	£6.25	£5.50	£4.78	£4.53	£3.05	£2.23
Supply of water total	£3.00	£2.67	£2.38	£2.38	£2.40	£1.46	£1.40
Sewerage services total	£3.36	£3.48	£3.05	£2.88	£2.53	£1.83	£1.55
Environment total	£3.67	£4.65	£4.85	£4.03	£4.26	£2.74	£1.72
Total	£17.27	£17.12	£15.57	£13.86	£13.69	£9.05	£6.90

Note: All starting 2020 bill levels, all per unit cost levels, all treatments.

Table 21 Average chosen bill increases, by household size

	Household size:				
	1 person	**2 people**	**3 people**	**4 people**	**5 + people**
Poor pressure	£0.03	£0.13	£0.09	£0.07	£0.11
Water quality, biological	£1.00	£1.48	£1.70	£1.97	£2.34
Water quality, aesthetic	£1.90	£3.20	£3.80	£5.11	£4.59
Unplanned interruptions	£0.19	£0.55	£0.35	£0.86	£0.43
Leakage	£1.07	£1.68	£1.61	£1.91	£1.65
Water restrictions	£0.09	£0.28	£0.25	£0.49	£0.50
Internal sewer flooding	£1.32	£1.88	£2.29	£3.01	£2.67
External sewer flooding	£0.28	£0.48	£0.63	£0.71	£0.70
Odour	£0.16	£0.27	£0.31	£0.36	£0.46
Bathing water	£0.31	£0.48	£0.55	£0.62	£0.63
River water	£1.73	£2.54	£2.47	£3.06	£3.16
Pollution incidents	£0.60	£0.87	£0.84	£1.15	£1.33
Land improved/conserved	£0.19	£0.29	£0.27	£0.33	£0.33
Water quality total	£2.93	£4.81	£5.60	£7.15	£7.04
Supply of water total	£1.35	£2.51	£2.21	£3.26	£2.58
Sewerage services total	£1.76	£2.63	£3.23	£4.09	£3.83
Environment total	£2.83	£4.18	£4.13	£5.16	£5.45
Total	£8.77	£14.07	£15.16	£19.29	£19.08

Note: All starting 2020 bill levels, all per unit cost levels, all treatments.

Table 22 Average chosen bill increases, by income

	Annual household income:					
	Under £10,000	**£10,000-£19,999**	**£20,000-£29,999**	**£30,000-£39,999**	**£40,000-£49,999**	**£50,000 or more**
Poor pressure	£0.08	£0.09	£0.10	£0.11	£0.12	£0.07
Water quality, biological	£0.91	£1.23	£1.46	£1.53	£1.54	£2.36
Water quality, aesthetic	£2.28	£2.40	£3.31	£3.47	£3.36	£5.41
Unplanned interruptions	£0.26	£0.27	£0.46	£0.41	£0.43	£0.95
Leakage	£0.99	£1.24	£1.48	£1.77	£1.55	£2.32
Water restrictions	£0.21	£0.33	£0.16	£0.30	£0.31	£0.32
Internal sewer flooding	£1.34	£1.64	£1.88	£2.08	£1.86	£3.17

External sewer flooding	£0.34	£0.41	£0.43	£0.56	£0.45	£0.75
Odour	£0.23	£0.27	£0.24	£0.28	£0.26	£0.38
Bathing water	£0.28	£0.34	£0.45	£0.57	£0.44	£0.81
River water	£1.43	£2.15	£2.22	£2.95	£2.25	£3.65
Pollution incidents	£0.64	£0.67	£0.72	£0.99	£0.79	£1.44
Land improved/conserved	£0.16	£0.24	£0.26	£0.29	£0.30	£0.42
Water quality total	£3.27	£3.72	£4.88	£5.11	£5.01	£7.85
Supply of water total	£1.45	£1.84	£2.10	£2.49	£2.29	£3.58
Sewerage services total	£1.91	£2.32	£2.54	£2.91	£2.58	£4.29
Environment total	£2.50	£3.39	£3.65	£4.79	£3.78	£6.33
Total	£9.14	£11.11	£13.17	£15.18	£13.66	£21.90

Note: All starting 2020 bill levels, all per unit cost levels, all treatments.

Table 23 Average chosen bill increases, by socio-economic group

	Socio-economic group:					
	A	**B**	**C1**	**C2**	**D**	**E**
Poor pressure	£0.12	£0.07	£0.09	£0.08	£0.16	£0.08
Water quality, biological	£1.70	£1.65	£1.47	£1.56	£1.40	£0.98
Water quality, aesthetic	£3.57	£3.36	£3.61	£3.24	£3.41	£2.09
Unplanned interruptions	£0.67	£0.44	£0.43	£0.46	£0.49	£0.35
Leakage	£1.93	£1.66	£1.43	£1.47	£1.25	£1.42
Water restrictions	£0.23	£0.27	£0.19	£0.41	£0.32	£0.33
Internal sewer flooding	£2.40	£2.00	£1.99	£1.91	£1.90	£1.49
External sewer flooding	£0.62	£0.49	£0.47	£0.53	£0.41	£0.39
Odour	£0.33	£0.25	£0.26	£0.35	£0.21	£0.23
Bathing water	£0.53	£0.53	£0.47	£0.40	£0.44	£0.36
River water	£2.55	£2.68	£2.47	£2.43	£2.02	£2.04
Pollution incidents	£1.03	£0.84	£0.88	£0.87	£0.70	£0.72
Land improved/conserved	£0.33	£0.29	£0.29	£0.25	£0.21	£0.21
Water quality total	£5.39	£5.09	£5.17	£4.87	£4.97	£3.14
Supply of water total	£2.83	£2.37	£2.05	£2.34	£2.06	£2.10
Sewerage services total	£3.35	£2.73	£2.72	£2.78	£2.51	£2.11
Environment total	£4.44	£4.34	£4.11	£3.95	£3.37	£3.32
Total	£15.97	£14.39	£14.05	£13.85	£12.50	£10.66

Note: All starting 2020 bill levels, all per unit cost levels, all treatments.

Econometric analysis

To augment the analysis presented in Table 19 above, we also undertook an econometric analysis regressing total chosen bill increases (the dependent variable) on age, household size, income and socio-economic group.

Coding variables for the econometric modelling

The variables representing age and income were each coded in two alternative ways:

- A continuous variables:

- For age, the variable is in years ranging from 19 to 89 years.
- For income, we take the mid-points of the bands offered in the survey question, with those who indicated an income of '£50,000 or more' assigned an income of £55,000 (thus the variable takes the values £5,000, £15,000, £25,000, £35,000, £45,000, and £55,000).

- As sets of dummies:
  - For age, seven dummies signifying whether the respondent is aged 18-29, 30-39, 40-49, 50-59, 60-69, 70-79, or 80-89 years
  - For income, six dummies signifying whether the respondent has a household income of under £10,000, £10,000-£19,999, £20,000-£29,999, £30,000-£39,999, £40,000-£49,999, or £50,000 or more.

Coding these as continuous variables provides results that are more parsimonious and straight-forward to interpret. On the other hand, the benefit of coding the variables as sets of dummies is that this allows the relationships between the explanatory characteristics and the dependent variable to be non-linear. For example, the estimated difference between those earning £10,000-£19,999 and those earning £20,000-£29,999 does not have to be the same as the estimated difference between those earning £20,000-£29,999 and those earning £30,000-£39,999. A separate regression coefficient is estimated for each group represented by a dummy, relative to a specified base group.

Household size is coded as a continuous variable (taking values 1 to 8, where 8 means '8 or more') and socio-economic group is coded as a set of dummies signifying whether the respondent is in group A, B, C1, C2, D or E.

Identifying variables for the model

A potential issue is that the explanatory characteristics of age, household size, income and socio-economic group may be highly correlated, meaning that the results of the regression analysis for one variable may depend on whether or not the other variables are included in the model. Therefore:

- As a first step we explore the correlations between these variables;
- Then, we follow the general to specific approach to specifying our econometric model, where we include all variables in a first run of the model, then remove any variables that are not statistically significant and re-run the model, and so on, until we reach a model that only includes statistically significant variables;
- Finally, we examine how the results for each variable change if we drop one of the other remaining variables in the final model.

This approach is followed using the continuous variables for age and income and, separately, using the dummies for these characteristics.

Correlations between explanatory variables

The correlation coefficients in Table 24 below show that the characteristics of age, household size, income and socio-economic group are indeed correlated to an extent. For example, being in the lowest socio-economic group (group E) is negatively correlated with income (as expected), with a correlation coefficient of -0.319. However, the results in this table do not indicate that there is too

much correlation between these variables to include them together as explanatory variables in a regression (i.e. the results to not suggest that we will have a problem with collinearity).

Table 24 Correlation coefficients

	Age	Household size	Income
Household size	-0.304	n/a	0.326
Income	-0.198	0.326	n/a
SE group A	0.075	0.056	0.243
"" B	0.034	0.004	0.224
"" C1	-0.122	-0.001	-0.007
"" C2	-0.006	0.024	-0.112
"" D	-0.026	0.031	-0.134
"" E	0.062	-0.116	-0.319

Note: All per unit cost levels, all starting 2020 bill levels, all treatments.

General to specific model specification

When all explanatory variables are included in the 'general' model there are statistically significant coefficients for age, household size and income, but not for socio-economic group – see Table 25. This is the case irrespective of whether age and income are included as continuous variables or as dummies. Note that when sets of dummies are used for age, income and socio-economic group, the youngest age group (18-29 years), lowest income level (under £10,000) and lowest socio-economic group (E) are omitted from the regression and therefore constitute the base.

The next step is to drop the socio-economic group variable(s) from the model, since this characteristic explains little variation in respondents' total chosen bill increases in the data (that is at least when the other characteristics are included in the model). Variables representing age, household size and income all remain statistically significant when socio-economic group is omitted, so the final model includes these three characteristics.

Table 25 Regression analysis of total chosen bill increases on age, household size, income, and socio-economic group – General model results

Model 1: Age and income as continuous variables			**Model 2: Age and income as dummies**		
Explanatory variable	**Coeff.**	**p-value**	**Explanatory variable**	**Coeff.**	**p-value**
Age (continuous)	-0.107***	0.001	Age: 30-39	-1.265	0.563
			Age: 40-49	-4.029*	0.065
			Age: 50-59	-3.779*	0.075
			Age: 60-69	-3.364	0.105
			Age: 70-79	-6.875***	0.002
			Age: 80-89	-9.675**	0.020
Household size	1.284***	0.001	Household size	1.386***	0.001
Income (continuous)	0.166***	0.000	Income: £10,000-£19,999	2.138	0.255
			Income: £20,000-£29,999	3.764*	0.051
			Income: £30,000-£39,999	5.007**	0.014
			Income: £40,000-£49,999	2.785	0.204
			Income: £50,000 or more	10.74***	0.000

SE group A	0.873	0.656	SE group A	0.878	0.658
SE group D	-0.256	0.885	SE group D	-0.124	0.946
SE group C1	0.180	0.916	SE group C1	0.225	0.898
SE group C2	1.042	0.580	SE group C2	1.218	0.530
SE group D	-1.022	0.632	SE group D	-0.898	0.680
Constant	11.71***	0.000	Constant	9.994***	0.000

Note: All per unit cost levels, all starting 2020 bill levels, all treatments. The continuous income variable is in 1000s of £s. Where dummies are included, the base groups are 18-29 years, income under £10,000, and socio-economic group (E). ***/**/* signifies statistical significance at 99%/95%/90% level.

Final regression results

The regression results in Table 26, which presents the results where age and income are coded as continuous variables, show that:

- **Older age** is associated with **lower chosen bill increases**; and
- **Larger household size** and **higher income** is associated with **higher chosen bill increases**.

These results still hold when each of the three variables are dropped in turn from the model.

Table 26 Regression analysis of total chosen bill increases on age, household size, and income – Final model results

	All variables included		Income dropped		Household size dropped		Age dropped	
	Coeff.	**p-value**	**Coeff.**	**p-value**	**Coeff.**	**p-value**	**Coeff.**	**p-value**
Age	-0.10***	0.001	-0.11***	0.000	-0.13***	0.000	n/a	n/a
Household size	1.29***	0.001	1.92***	0.000	n/a	n/a	1.62***	0.000
Income	0.17***	0.000	n/a	n/a	0.20***	0.000	0.18***	0.000
Constant	11.7***	0.000	15.5***	0.000	15.4***	0.000	4.94***	0.000

Note: All per unit cost levels, all starting 2020 bill levels, all treatments. The continuous income variable is in 1000s of £s. ***/**/* signifies statistical significance at 99%/95%/90% level.

These overall findings do not change when age and income are coded as dummies, as can be seen from Table 27.[9]

Regarding age, these results show that the largest estimated effects are those for the two oldest age groups (70-79 and 80-89 years), confirming the results in Table 19 above. Larger household size is associated with higher chosen bill increases. As regards income, these the largest estimated effect is that for the highest income group (£50,000 or more).

Excluding the variables for each characteristic in turn does not have a major impact of the results for the remaining variables.[10]

[9] Note that when sets of dummies are used for age, income and socio-economic group, the youngest age group (18-29 years), lowest income level (under £10,000) and lowest socio-economic group (E) are omitted from the regression and therefore constitute the base.

[10] There are some minor changes: when income is dropped from the model, the results for the 40-49 and 50-59 years age bands become insignificant; and when household size is dropped, the results for the 60-69 years age band and the £40,000-£49,999 income band become significant.

Table 27 Regression analysis of total chosen bill increases on age, household size, and income – Final model results, age and income as dummies

	All variables included		Income dropped		Household size dropped		Age dropped	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Age:								
30-39	-1.35	0.537	0.11	0.956	-1.44	0.509	n/a	n/a
40-49	-4.02*	0.065	-1.97	0.344	-3.90*	0.074	n/a	n/a
50-59	-3.83*	0.071	-2.79	0.166	-4.12*	0.052	n/a	n/a
60-69	-3.36	0.103	-1.96	0.318	-4.30**	0.036	n/a	n/a
70-79	-6.81***	0.002	-6.28***	0.003	-7.99***	0.000	n/a	n/a
80-89	-9.52**	0.021	-8.02**	0.043	-10.9***	0.008	n/a	n/a
Household size	1.40***	0.000	1.97***	0.000	n/a	n/a	1.67***	0.000
Income:								
£10,000-£19,999	2.20	0.230	n/a	n/a	2.52	0.167	1.74	0.340
£20,000-£29,999	3.83**	0.035	n/a	n/a	4.36**	0.016	3.52**	0.052
£30,000-£39,999	5.15***	0.007	n/a	n/a	5.98***	0.002	5.05***	0.008
£40,000-£49,999	2.87	0.158	n/a	n/a	4.11**	0.041	2.86	0.160
£50,000 or more	10.9***	0.000	n/a	n/a	12.3***	0.000	10.9***	0.000
Constant	10.1***	0.000	11.7***	0.000	13.2***	0.000	6.05***	0.000

Note: All per unit cost levels, all starting 2020 bill levels, all treatments. The base groups are 18-29 years, income under £10,000, and socio-economic group (E). ***/**/* signifies statistical significance at 99%/95%/90% level.

11.6 Results for particular vulnerable groups

For the purposes of our analysis we defined eight different vulnerability characteristics:

- Aged 75+: Respondent is aged 75 or over
- Worry about bill (agrees that *"I worry about not being able to afford my water bill"*)
- Can't afford bill (agrees that *"I already can't afford my water bill"*)
- Receive help to pay bill
- English not first language
- Disabled person (at least one disabled person in the household)
- Receive/received benefits (receive any type of benefits, or received benefits in the past)
- Low income (household income below £10,000, or household income below £20,000 for households of 4 or more people)

In addition, we also define three combined groups comprising of those who are vulnerable on *any one* of a set of (similar) vulnerability characteristics:

- Financial: Any of worry about bill, can't afford bill, receive help to pay bill, receive/received benefits, or income below £10,000.
- Physical: Either of aged 75+ or disabled person.
- Any reason: Any of the eight vulnerability characteristics.

Average chosen bill increases among those with these vulnerability characteristics vary either side of the overall mean, although in general they are lower than average bill increases chosen by those without any of these characteristics (+£14.66), as shown in Table 28 overleaf.

Furthermore, it should be noted that the differences between those who are vulnerable and those who are not vulnerable are often not statistically significant. In particular, the differences (compared to those who are not vulnerable on the relevant characteristic) in terms of the total bill increases chosen are *not statistically significant* for those who worry about bill their bill, who can't afford their bill, whose first language is not English, and who have a disabled person in their household.

Conversely, the differences between those who are vulnerable and those who are not vulnerable (in terms of total bill increases chosen) *are statistically significant* for the vulnerability characteristics of age, receive/received benefits and income below £10,000, as well for the combined groups of those with financial vulnerability and those who are vulnerable for any reason.

Average chosen bill increases are lowest for those aged above 75 (+£8.55) and for those with annual incomes below £10,000 (+£10.74).

Table 28 Average chosen bill increases, by vulnerability characteristic

	Vulnerability characteristic:											
	Aged 75+	**Worry about bill**	**Can't afford bill**	**Receive help to pay bill**	**English not first language**	**Disabled person**	**Receive/ received benefits**	**Low income**	**Financial**	**Physical**	**Any reason**	**None (Not vulnerable)**
Poor pressure	£0.04	£0.05	£0.10	£0.18	£0.11	£0.11	£0.09	£0.07	£0.09	£0.07	£0.10	£0.10
Water quality, biological	£0.90	£1.46	£1.93	£1.53	£1.70	£1.42	£1.33	£1.14	£1.43	£1.42	£1.36	£1.57
Water quality, aesthetic	£2.34	£2.68	£3.42	£4.32	£4.04	£3.00	£3.01	£2.34	£3.23	£3.08	£2.97	£3.35
Unplanned interruptions	£0.11	£0.30	£0.23	£0.84	£0.27	£0.43	£0.44	£0.39	£0.40	£0.38	£0.39	£0.55
Leakage	£1.16	£1.43	£1.38	£1.46	£1.23	£1.54	£1.27	£1.06	£1.45	£1.40	£1.47	£1.69
Water restrictions	£0.19	£0.32	£0.54	£0.95	£0.45	£0.38	£0.33	£0.40	£0.31	£0.31	£0.35	£0.23
Internal sewer flooding	£1.23	£1.84	£2.18	£1.59	£2.05	£1.93	£1.84	£1.69	£1.92	£1.90	£1.81	£2.05
External sewer flooding	£0.25	£0.45	£0.53	£0.63	£0.58	£0.47	£0.46	£0.43	£0.48	£0.48	£0.45	£0.51
Odour	£0.15	£0.28	£0.29	£0.49	£0.30	£0.27	£0.27	£0.28	£0.27	£0.28	£0.26	£0.27
Bathing water	£0.22	£0.43	£0.45	£0.31	£0.46	£0.44	£0.37	£0.32	£0.43	£0.42	£0.40	£0.52
River water	£1.48	£2.35	£2.50	£1.49	£2.62	£2.36	£1.88	£1.80	£2.25	£2.20	£2.25	£2.66
Pollution incidents	£0.36	£0.84	£0.99	£0.56	£0.98	£0.81	£0.67	£0.76	£0.77	£0.78	£0.76	£0.96
Land improved/conserved	£0.13	£0.27	£0.23	£0.21	£0.18	£0.24	£0.20	£0.19	£0.23	£0.24	£0.23	£0.32
Water quality total	£3.28	£4.19	£5.45	£6.03	£5.85	£4.52	£4.43	£3.55	£4.74	£4.57	£4.43	£5.03
Supply of water total	£1.45	£2.05	£2.15	£3.26	£1.95	£2.36	£2.04	£1.85	£2.15	£2.09	£2.22	£2.47
Sewerage services total	£1.63	£2.57	£3.01	£2.71	£2.93	£2.67	£2.57	£2.40	£2.67	£2.66	£2.51	£2.82
Environment total	£2.19	£3.89	£4.16	£2.56	£4.25	£3.85	£3.12	£3.07	£3.68	£3.64	£3.64	£4.46
Total	£8.55	£12.76	£14.53	£14.56	£14.98	£13.39	£12.09	£10.74	£13.19	£12.88	£12.78	£14.66

Note: All starting 2020 bill levels, all per unit cost levels, all treatments.

11.7 Results for the alternative treatments

Disposable income treatment

At the aggregate level, the disposable income treatment had no statistically significant impact. This suggests that our aggregate baseline treatment results are robust to the presentation of disposable income information. As can be seen from Table 29, the differences in chosen bill increases per attribute between the baseline and disposable income treatments are relatively small, ranging from -£0.35 (for aesthetic water quality) to +£0.35 (for improvement of rivers). The difference between these two treatments in terms of the total chosen bill increase is just +£0.19.

However, we have also explored this treatment effect separately for the various different vulnerable groups and found that there is an effect for those who are vulnerable due to disability. For this subgroup the total chosen bill increase was, on average, £4.75 per year higher if their remaining disposable income was shown. This may be because, for this especially vulnerable group, showing their remaining disposable income may in fact reassure them that their water bill will not in fact consume a very large share of their disposable income. This higher total chosen bill increase for this group under the disposable income treatment was driven in particular by higher chosen bill increases for internal and external flooding and the environment attributes.

Table 29 Impacts of presenting individual's disposable incomes on chosen bill increases

	Chosen bill increases under each treatment		
	Baseline Treatment	DI Treatment	Treatment effect
Poor pressure	£0.11	£0.09	-£0.02
Water quality, biological	£1.39	£1.41	+£0.01
Water quality, aesthetic	£3.56	£3.21	-£0.35
Unplanned interruptions	£0.54	£0.43	-£0.12
Leakage	£1.58	£1.57	-£0.02
Water restrictions	£0.24	£0.30	+£0.06
Internal sewer flooding	£1.97	£2.11	+£0.14
External sewer flooding	£0.52	£0.50	-£0.02
Odour	£0.29	£0.30	+£0.01
Bathing water	£0.46	£0.53	+£0.08
River water	£2.23	£2.58	+£0.35
Pollution incidents	£0.84	£0.89	+£0.04
Land improved/conserved	£0.27	£0.28	+£0.01
Total bill change	£14.02	£14.20	+£0.19

Comparative industry positions treatment

The comparative industry positions treatment was applied for four attributes in particular in the experiment (i.e. the four for which relevant comparable industry average information is available):

- Unplanned interruptions
- Internal sewer flooding
- Leakage
- Pollution incidents

For these attributes, under the comparative industry positions treatment information on the industry average service level was presented alongside the sliders. An informative measure to examine in order to assess the impact of this treatment is the share of respondents who selected service levels close to (within +/- 1%) the industry average for these four attributes.

The results reveal a greater tendency for participants' chosen service levels to cluster around the industry average when this information was shown, as can be seen from Table 30 below. For example, for pollution incidents the share of respondents who chose a service level within +/- 1% of the industry average was 6.7% under the comparative industry positions treatment, compared to just 1.6% under the baseline treatment. Thus, presenting this information did have an impact on choices. However, the shares who chose service levels close to the industry average are not especially large under either treatment, and the treatments effects (although statistically significant) are not large enough to have resulted in statically significant differences in the average chosen bill increases for these attributes.

These results imply that it is important to interpret the bill impact results in the context of whether or not this anchoring effect is present (i.e. whether or not the comparative industry information was presented). The general approach taken in the SP surveys in WP1 and WP2 was to provide comparative industry information and to instruct respondents to take note of this information. However, the dynamics of the choice experiment approach in WP1 and WP2 mean that it is less likely that this anchoring effect is present, since respondents make choices between options with specified service levels and prices rather than choosing a service level for each attribute individually.

Table 30 Impacts of presenting individual's disposable incomes on chosen bill increases

	Shares within +/- 1% of industry average:		
	Baseline Treatment	**CP Treatment**	**Treatment effect**
Unplanned interruptions	1.2%	4.4%	+3.2**
Internal sewer flooding	1.6%	4.8%	+3.1**
Leakage	4.0%	9.5%	+5.5***
Pollution incidents	1.6%	6.7%	+5.1***

Note: All per unit cost levels, all starting 2020 bill levels. ***/**/* signifies statistical significance at 99%/95%/90% level.

Reframing the service levels as frequencies rather than quantities

This treatment was applied for eight attributes:

- Poor pressure
- Water quality, biological
- Water quality, aesthetic
- Unplanned interruptions
- Internal sewer flooding
- External sewer flooding
- Odour
- Pollution Incidents

Under this treatment the service levels for these attributes were presented as frequencies instead of quantities. For example, for internal sewer flooding under the baseline treatment the service level was given as the number of incidents per year (e.g. "Incidents of sewer flooding of living areas inside

properties per year: 1,919"), whereas under the alternative treatment it was given as a frequency in terms of hours and minutes (e.g. "One incident of sewage flooding of living areas inside a property every 4hrs 34min").

The first impact of this treatment that should be highlighted is that reframing the service levels as frequencies made respondents less able to respond in a realistic way. Under the baseline treatment 20% said they felt unable to make their choices in a realistic way, whereas under the reframed levels treatment this share was 27%. This seven percentage point difference is statistically significant at the 99% percent level.

This implies that individuals were less able to understand the service levels when these were presented as frequencies. This is an important finding in the context of the wider research programme, since the stated preference survey in Work Package 1 used the same presentation of the service levels as in our baseline treatment, which was better understood. Thus this result from Work Package 5 supports the choice of service level presentation used in Work Package 1.

As can be seen from Table 31 below, reframing the service levels in this way had a statically significant impact on the *service levels chosen* for a number of different attributes. In particular, a worse level of service was typically chosen when the levels were reframed in this way (note that for these attributes, except for biological water quality, higher values for the service level imply worse service).

Table 31 Impacts of presenting individual's disposable incomes on chosen bill increases

	Chosen service levels under each treatment:		
	Baseline Treatment	**Reframed levels**	**Treatment effect**
Poor pressure	14.1	15.4	+1.3***
Water quality, biological	9996.3	9996.3	0.0
Water quality, aesthetic	5595	5812	+216***
Unplanned interruptions	38686	39777	+1092***
Internal sewer flooding	1774	1808	+33*
External sewer flooding	9991	10134	+143
Odour	5392	5515	+123*
Pollution incidents	184.2	191.0	+6.7**

Note: All per unit cost levels, all starting 2020 bill levels. ***/**/* signifies statistical significance at 99%/95%/90% level.

However, the service level treatment effects are not matched by statistically significant chosen *bill changes* from the reframing of the service level attributes – see Table 32 below. This discrepancy between the treatment effects for the chosen bill changes and chosen service levels is possible due to the fact that, in the experiment set up, the service levels and bill increases were not perfectly correlated[11] (e.g. the correlation between the service levels and bill increases for the water pressure attribute, across all starting 2020 bill levels and all unit cost settings, was 0.81).

Table 32 Impacts of reframing service levels as frequencies on chosen bill increases

	Chosen bill increases under each treatment	

[11] Which was due to variation in starting 2020 bill levels across respondents (with consequent variation in bill impacts per unit change in service), variation in the per unit costs associated with service level changes (depending on the setting the respondent was assigned to), and differences in the per unit bill impacts for service levels above and below 2020 status quo service level.

	Baseline Treatment	Reframed levels	Treatment effect
Poor pressure	£0.11	£0.08	-£0.04
Water quality, biological	£1.39	£1.81	£0.42**
Water quality, aesthetic	£3.56	£3.08	-£0.48
Unplanned interruptions	£0.54	£0.34	-£0.20*
Internal sewer flooding	£1.97	£1.82	-£0.15
External sewer flooding	£0.52	£0.47	-£0.05
Odour	£0.29	£0.24	-£0.06
Pollution incidents	£0.84	£0.79	-£0.05

Note: All per unit cost levels, all starting 2020 bill levels. ***/**/* signifies statistical significance at 99%/95%/90% level.

LE London Economics

Somerset House, New Wing, Strand,
London, WC2R 1LA, United Kingdom
info@londoneconomics.co.uk
londoneconomics.co.uk
@LondonEconomics
+44 (0)20 3701 7700