Working Paper No. 655

March 2020

Eliciting trade-offs between water charges and service benefits in Scotland

Cameron A. Belton[a] , Ciarán Lavin[a] and Peter D. Lunn[a,b]*

Abstract: If it is the responsibility of a regulatory body to decide where to prioritise future investment, then it is important to understand the priorities of the citizenry it represents. This paper, in collaboration with the OECD and the Scottish water industry, presents the results of an online (n= 500) and face-to-face laboratory (n= 99) study that utilised experimental behavioural science to explore how Scottish citizens trade-off costs and potential improvements to their water service. Participants' priorities for investment were elicited using a novel 'slider task' methodology that forced them to explicitly consider the trade-offs required to allocate limited resources across multiple possible water service improvements. The provision of additional cost and timing information was systematically varied. Results suggest that citizens are increasingly accepting of price rises when provided this information. Results also suggest that citizens' priorities for specific improvements are not sensitive to the costs of different improvements but are sensitive to the lengths of time improvements take to be made. Findings from this study are designed to inform the regulatory process of the Scottish water industry and highlight the potential role of behavioural science in regulation more generally.

**Corresponding Author: cameron.belton@esri.ie*

Acknowledgements: This work was funded by the Organisation for Economic Co-operation and Development (OECD) under a framework contract designed to investigate alternative approaches to implementing regulatory policy using behavioural insights. For helpful discussions we thank Faisal Naru, Filippo Cavassini, Lorenzo Casullo and Alex Durand of the OECD, Alan Sutherland, Kerstin McEwan and David Satti of the Water Industry Commission for Scotland, and members of the Customer Forum and Scottish Water. We also thank attendees of seminars at the Economic and Social Research Institute, the 2019 IAREP-SABE Conference, and the 12th Annual Economic and Psychology Conference, and the other members of the Behavioural Research Unit, for feedback on earlier versions of this work.

a Economic and Social Research Institute, Dublin
b Department of Economics, Trinity College Dublin

ESRI working papers represent un-refereed work-in-progress by researchers who are solely responsible for the content and any views expressed therein. Any comments on these papers will be welcome and should be sent to the author(s) by email. Papers may be downloaded for personal use only.

Non-Technical Summary

This study tested how giving households details about the costs and benefits of improvements to water services alters their preferences for what specific service improvements they think money should be spent on. The study was a computerised experiment undertaken by 599 Scottish citizens. Of these, 500 completed the experiment online, 99 face-to-face.

The experimental approach makes it possible to measure baseline responses, then to observe how these responses change when participants are given specific additional pieces of information. The study tested for and measured any systematic influence associated with the communication of specific costs and/or benefits.

Eliciting Baseline Responses

Research Question: *What are citizens' general preferences for different specific service improvements?*

Research Question: *What would citizens deem to be an acceptable price rise for water charges to invest in specific service improvements?*

The study first recorded baseline preferences for specific service improvements. Three initial measures were taken. First, participants were given descriptions of seven aspects of water service that the Scottish water industry might want to improve - reducing interruptions to supply, reducing external sewer flooding, connecting rural supplies to the water system, increasing the share of renewable energy generated by the water industry, reducing leakage in supply pipes, improving drinking water look, taste and smell, and improving cleanliness of rivers, seas and beaches. Participants then ranked these in order of importance to them. Four of these were *target improvements* (which, unbeknownst to participants, were the four improvements that were pre-selected to be used in all later stages). These were broadly *personal* (reducing interruptions to supply, reducing external sewer flooding), or *societal* (connecting rural supplies to the water system, increasing the share of renewable energy generated by the water industry), improvements, and were selected by the researchers in agreement with the water industry.

Second, participants were asked to allocate an amount of investment between the four *target improvements*, using a set of sliders that required them to split 100% of investment between the four improvements. The design of these sliders made the trade-offs required salient: in order to invest more in one improvement, the amount spent on others had to be reduced.

Lastly, participants were asked the maximum they would be willing to increase their current water charge by if the revenue raised was to be spent on the *target improvements*. Each participant was told their current annual water charge (based on the council tax band they provided). Participants could increase this amount, or leave it at its current level, but could not decrease their water charge below its current level. Increases were in expressed in pounds and pence with no reference to inflation.

Overall, at the individual level, there was a lot of variety in preferences: different participants had different priorities for improvements, both when ranking them and when allocating investment using the sliders. This largely averaged out, meaning that no single improvement clearly stood out as most (or least) important. There was strong consistency between rankings of improvements and allocations using the sliders. Younger participants and those who undertook the study face-to-face were more likely to favour *societal* improvements.

Approximately 60% of participants indicated that they would accept some price rise to invest in the *target improvements*. Of these, the average acceptable price rise was £18.41 (a 4.2% increase on current charges). The remaining 40% indicated that they would not accept any price rise. Face-to-face participants were willing to increase their charges by approximately double that of online participants.

Answer: Individual participants displayed strong and consistent preferences. However, because households had contrasting views, across the population as a whole no given service improvement stood out as more or less popular.

Answer: Most citizens were willing to increase their water bill when the revenue was to be used to improve specific service benefits. These citizens were willing to increase their water bills by 4.2%.

Effect of Learning about Costs and Time Horizons

Research Question: *How does providing people with the costs and time horizons of different specific service improvements alter their priorities and acceptance of price rises to fund improvements?*

Revealing more detailed information about the cost and timing of *target improvements* – what people can expect in return for their money – could systematically change participants' views about where investment should be directed and what constitutes an acceptable charge. Having established their baseline responses, we showed participants the estimated unit costs of making each improvement (based on approximations provided by the Scottish water industry). This detailed the amount of the improvement expected for a given increase in investment. It was new information that participants would be very unlikely to know beforehand. They were then asked again to allocate investment to each *target improvement* using sliders. They could also change the price rise that they deemed to be acceptable for these improvements.

Next, they were shown different time horizons for the benefits of each improvement to be felt. These were not based on industry estimates but were plausible round numbers. Time horizons were pseudo-randomised between participants: the *personal* improvements were randomly allocated "Immediately" or "5 Year" time horizons, and the *societal* improvements were randomly allocated "10 Year" or "25 Year" time horizons. As before, they were then asked to allocate investment to each *target improvement* using the sliders and could change the acceptable price rise as they wished.

Thus, by incrementally providing information this study measured if, and how, learning additional information alters people's views about where investment should be directed and how much they might pay for it.

The results showed that, in fact, learning about costs and time horizons had little impact on citizens' stated priorities. The most notable change was that the allocations to improvements that would be realised immediately, rather than in 5-25 years' time, increased by 1.4 percentage points. However, while this is consistent with previous findings that show that people prefer immediate gains to later ones, the effect was much smaller than is generally observed. Respondents were largely willing to take a long-term view, at least when investment was directed to their preferred improvements.

Changes to acceptable charges were also minor. When costs were revealed, 14.7% increased their acceptable price rise, and 1.2% reduced it (84.1% of participants did not change it), increasing the average acceptable price rise by £1.06. Likewise, when time horizons were revealed, 7.5% increased their acceptable price rise, and 1.3% reduced it (91.2% of participants did not change it), increasing the average acceptable price rise by a further £0.47.

Answer: Priorities for improvements are hardly changed by learning about the costs and time horizons involved.

Answer: Information about costs and time horizons increases the acceptability of price rises for some participants.

Summary

The experimental method provided clear answers to the research questions asked. Overall, the findings can be summarised as four key results:

1. There is substantial variation in the preferences of participants regarding what improvements to water services they would most like to see, with no clear overall preference for personal or societal benefits.
2. Participants, in general, are willing to accept modest additional water charges to improve specific services.
3. People's priorities for improvements are not changed substantially by learning about the costs and time horizons involved.
4. Learning about costs and time horizons led some participants to deem a further small price rise acceptable.

1. Introduction

Ideally, regulation of the water industry in Scotland will be informed by the best possible evidence regarding what Scottish people want from their water services. Yet how can we understand these preferences? The industry has been described as a 'silent service' (WICS, 2017). Many people may have little knowledge of straightforward elements of their water system, such as the chances of service interruptions, or of more complex industry-level issues, such as the need for future investment. It is therefore possible that people's views will be sensitive to the extent of their knowledge, so that as they learn more about the system their preferences change. This study presents a systematic experimental approach to this issue, focused on how the provision of information about the costs and benefits of improvements to water services affects households' preferences regarding investment priorities.

It is inevitably a challenge for any large utility to decide where to prioritise future investment. Citizens in Scotland do appear willing to engage with the water industry on these matters, and previous qualitative studies have demonstrated diverse opinions regarding the desired direction of future investment (e.g. Walker, 2017). However, it is very unlikely that households have a set of well-formed, stable preferences over water charges and systems; rather they "construct" their preferences as information and decisions are put in front of them. Controlled behavioural experiments are a good tool for investigating these issues, because the level and type of information provision can be experimentally controlled and tested. For example, experimental findings suggest that preferences for environmental services can differ when these are elicited as either choices between options or as independent valuations (e.g. Irwin *et al*., 1991).

This is the second in a series of experimental research studies designed to illuminate the views of Scottish citizens regarding water. The first study (Belton *et al*., 2020) elicited attitudes to potential changes in water charges. The results revealed some tolerance for modest price rises in the short and medium term, as well as clear preferences for the trajectory of price changes across multiple years. However, these responses were obtained in the absence of any information about how the money might be spent or over what timescale any benefits might accrue. In principle, people's willingness to pay might be altered substantially by knowing what they might get in return for their money. Sensitivity of preferences to information on costs and benefits of service improvements might arise for several reasons. People may be overly optimistic or pessimistic about the amount or speed of improvement associated with a given increase in charges, meaning that relevant information alters whether they view the return as worth the increase. This could apply differentially across the range of potential improvements, with the potential for systematic overestimation or underestimation of improvements of a particular type. Willingness to pay may also be contingent on the money being used to address the specific improvements that an individual would like to see prioritised, which may vary substantially between different individuals.

A primary aim of this study was to record Scottish citizen's preferences when the trade-off between costs and benefits of additional investment were made explicit and, hence, respondents were forced to balance different improvements to the system both against each other and against their costs. An increase in investment for any given improvement may come from two possible strategies. Either the absolute level of investment within the industry must rise, or the share of investment directed towards the intended improvement must increase. In the latter case, investment must be taken away from another potential improvement.

This study required a sample of Scottish citizens to consider these very sorts of trade-offs. Through a sequence of experimental stages, participants indicated which types of water service improvements they thought should be prioritised over others, and by how much. The available information about costs and timescales was carefully controlled at each stage and the experimental design allowed them to express their preferences using an interface via which they could vary both overall investment and its allocation across potential improvements.

Previous survey research has provided some empirical insights into how individuals rank benefits in the Scottish water industry (Scottish Water, 2017), but these were elicited in the absence of information about what benefit might be obtained for a given increase or decrease in charges. By using the controlled behavioural experiments described here, the present study obtained measures of more informed household preferences, asking what people want when told what they can expect in return for their money, and when they can expect to see these benefits felt. There were four specific research questions:

1) What are citizens' general preferences for different specific service improvements?
2) What would citizens deem to be an acceptable price rise for water charges to invest in specific service improvements?
3) How does providing the costs of different specific service improvements alter the priorities and willingness-to-pay towards improvements of citizens?
4) How does providing the time horizons of different specific service improvements alter the priorities and willingness-to-pay towards improvements of citizens?

In order to answer these specific research questions, this study implemented five distinct experimental stages. In addition, two more general research questions were addressed across multiple stages: (i) Is there a general preference for service benefits likely to affect the individual household directly (e.g., reduced likelihood of service interruption) as opposed to benefits accruing to society more generally (e.g., increasing the share of renewable energy)? (ii) Are there strong differences in responses across individuals? As there were theoretical, methodological and practical differences between each experimental stage, we report the motivation, design and results of each stage sequentially, in the order in which participants completed them. A general discussion of the main overall findings, and the potential policy implications of these, then follows. First, we outline the general design of the overall study.

2. General Study Design

This study was a computerised experiment that presented participants with a variety of tasks, aimed at eliciting different measures of their attitudes towards investment for improvements to the water sector in Scotland. As well as presenting participants with multiple tasks, the information about possible investment and improvements varied across different tasks, which all participants undertook (a within-subject design). In addition, for some aspects of the experiment, the exact information that participants saw differed across participants (a between-subject design).

A total of 599 Scottish citizens took part in this study. 500 completed it online, and 99 completed it face-to-face in a mobile laboratory. In the online study, participants who were already signed up to an online market research company were invited to follow an email link on their personal computers to begin the experiment. The experiment itself took approximately 20 minutes to complete, and participants received the industry standard participation fee determined by the market research company (approximately £2.00). In the face-to-face study, participants were recruited by a market research company to attend four locations in Scotland: Glasgow, Inverness, Aberdeen and Edinburgh. Participants attended the session for approximately 30 minutes in total and were paid £20 (approximately the standard fee for face-to-face studies).

The studies were identical across both platform types. The studies used the experimental platform Gorilla Experiment Builder (www.gorilla.sc). All instructions were identical, but were read aloud by an instructor in the face-to-face studies. Efforts were made to match the demographic characteristics of participants across both platform types, and details of this can be found in Appendix H. Systematic differences in responses based on platform type would raise important methodological questions about the reliability of different study platforms in eliciting truthful or consistent responses in studies such as this. A comparison of results by platform type can be found in Appendix H.

3. Stage 1 – Ranking Task

Research Question: *What are citizens' general preferences for different specific service improvements?*

3.1. Introduction

In the first stage, participants were provided with descriptions of seven possible water service issues that the Scottish water industry might want to prioritise investment towards to improve. The aim of this stage was to ascertain an initial sense of citizens' priorities towards a wide array of potential service improvements that the Scottish water industry might choose to invest in. For each of the seven potential service issues, participants were given a description of the potential issue, as well as a potential improvement to the issue. Details of the descriptions provided for each issue can be found in Appendix A. It is important to note that, while the description of each service improvement was designed with input from the Scottish water industry, this was not a qualitative study that aimed to provide participants with detailed information about these issues. It remains possible that framing these descriptions in a different way might alter responses. The main aim of this stage was to generate a general inference of citizens' attitudes to different water service issues, and to provide a benchmark from which comparisons could be made when these preferences were elicited again using different methodologies and in the presence of additional information about these service issues.

Four of these potential improvements used in this stage were *target improvements* – which were to be used in the later stages of the experiment:

1. Reducing the number of short term interruptions to supply (Interruptions to Supply)
2. Reducing the number of external sewer flooding incidents (External Sewer Flooding)

3. Connecting rural supplies to the water system (Rural Supplies)[1]
4. Increasing the share of renewable energy generated by the water industry (Renewable Energy)

These four improvements were determined as most appropriate for two reasons. The first is that, whilst participants in this study were responding to specific improvements in the water industry, these decisions could be indicative of more general preferences of Scottish citizens. Two of these improvements were broadly *personal* improvements, where the improvement would directly affect water in or around the home, namely reducing interruptions to supply and external sewer flooding. The remaining two were more *societal* improvements, where for the vast majority of citizens, the benefits would not have a direct positive impact on them. Connecting rural supplies would not directly affect the vast majority of citizens. Increasing the proportion of renewable energy generated by the water industry would not have direct benefit to citizens either. Systematic preferences for one category of improvements over the other might indicate the level of pro-social attitudes of Scottish citizens with regards to the water industry. The second reason was that, for later stages in the study, approximations of real cost estimates for improvements were given to participants. The Scottish water industry was able to provide more confident estimates of these specific improvements than other potential improvements, making them more suitable for experimental manipulation.

The remaining three improvements in this stage were other plausible water service improvements for the Scottish water industry:

1. Reducing leakage in supply pipes (Supply Pipe Leakages)
2. Improving drinking water look, taste and smell (Drinking Water Quality)
3. Improving cleanliness of rivers, seas and beaches (Beach/ River Cleanliness)

At this stage, participants were not aware which subset of improvements were defined as *target* improvements and would be used in the later stages of the study. After taking time to read the descriptions of each issue and its improvement participants were asked to rank the seven improvements in order of importance to them (where 1 = most important and 7 = least important). By asking participants to rank seven possible improvements, the results from this stage enabled a measure of the following:

a) The absolute ranking of each *target improvement* relative to one another

[1] A small number (approximately 4%) of households in Scotland are not connected to the public water supply, but are instead connected to a private supply. These are predominantly found in rural areas.

b) How much participants care about the *target improvements* in general (relative to other improvements)

3.2. Results

The order in which the service descriptions and improvements were listed was randomised across participants. There was strong evidence of an order effect, where descriptions/ improvements viewed earlier were ranked as more important. As the order was randomised at the individual level this should not adversely affect the findings in later stages, because no one service improvement should have disproportionately benefitted from these order effects, but this is a potentially important methodological finding. Further details may be found in Appendix I.

Table 1 below reports the average ranking of each of the seven possible improvements from most important to least important, as well as the likelihood any given improvement was to be ranked as 1^{st}, 7^{th} or 1^{st} – 4^{th} most important, by improvement and improvement type. It is also reports the rankings, and likelihood to be ranked 1^{st} or 4^{th}, of the four *target improvements* relative to one another.

In general, there does not appear to be strong evidence of a systematically and substantially most (or least) important improvement. Overall, Beach/ River Cleanliness was perceived to be the most important improvement (3.24/7), as indicated by a lower average ranking score, and Interruptions to Supply least important (4.79/7). Appendix B reports pairwise tests of the medians of differences for each pair of improvements. This analyses reveals that participants were not significantly more likely to rank Beach/ River Cleanliness as more important than Supply Pipe Leakages in a pairwise comparison (p= 0.248), and nor were they significantly more likely to rank Interruptions to Supply as less important than Rural Supplies (p= 0.150). Drinking Water Quality was ranked "most important" (i.e. 1^{st}) most frequently (20.2%), and Interruptions to Supply least frequently (7.3%). Renewable Energy was ranked "least important" (i.e. 7^{th}) most frequently (23.5%) and Beach/ River Cleanliness least frequently (5.2%).

On average, when pooling rankings by *target improvements* (4.40/7) and other improvements (3.47/7), the *target improvements* were ranked as less important. In addition, on average, any given *target improvement* was almost half as likely to be ranked as "most important" as the other improvements. Likewise, any given *target improvement* was ranked as one of the top four most important improvements almost 50% of the time on average. The same was true almost

70% of the time for any given other improvement on average. Among the four *target improvements*, External Sewer Flooding was ranked most important on average (2.11/4) and was ranked "most important" most frequently (for 33.9% of participants).

	Overall				Target Improvements		
	Average Ranking (/ 7)	**Proportion of Ranking 1st**	**7th**	**1st - 4th**	**Average Ranking (/ 4)**	**Proportion of Ranking 1st**	**4th**
Beach/ River Cleanliness	3.24	19.2%	5.2%	73.0%	-------	-------	-------
Supply Pipe Leakages	3.38	19.4%	5.5%	71.3%	-------	-------	-------
External Sewer Flooding	3.74	12.9%	7.3%	65.1%	2.11	33.9%	10.4%
Drinking Water Quality	3.80	20.2%	15.9%	60.3%	-------	-------	-------
Renewable Energy	4.41	13.2%	23.5%	48.7%	2.50	28.4%	29.2%
Rural Supplies	4.64	7.8%	20.5%	42.4%	2.64	18.9%	27.6%
Interruptions to Supply	4.79	7.3%	22.0%	39.2%	2.74	18.9%	32.9%
Target Improvements	4.40	10.3%	18.4%	48.9%	-------	-------	-------
Other Improvements	3.47	19.6%	8.9%	68.2%	-------	-------	-------

Table 1. Average ranking and likelihood to be ranked as 1st, 7th or 1st - 4th most important by improvement, from most important to least important (*target improvements* highlighted in **bold**)

Histograms of the distributions of rankings can be found in Appendix C. Analysis of these distributions reveals evidence of a 'U-shaped' distribution for Renewable Energy and Drinking Water Quality – suggesting that these were more divisive improvements, as participants disproportionately ranked them as most or least important. Tests of equality of variance across improvements, found in Appendix C, offers some evidence in support of the claim that some improvements were more divisive than others.

Because the importance each individual gave to the *target improvements* might have influenced their responses in the later stages of the study, we used the responses from Stage 1 to construct a relevant measure. Since the same four improvements were used as the *target improvements* for all participants, it is possible to sum the number of these *target improvements* that were ranked 1st – 4th most important for each participant. This provides a metric to be used as a proxy for the relative importance of these *target improvements* for each participant among the wider set of potential improvements. Table 2 reports the proportion of participants categorised by this importance metric: the number of *target improvements* that were ranked 1st – 4th most important.

Number of Target Improvements Ranked 1st – 4th	**1**	**2**	**3**	**4**
Number of Participants	136	355	107	1
Percentage of Participants	22.7%	59.3%	17.9%	0.2%

Table 2. Proportion of participants by number of *target improvements* ranked 1st – 4th most important

It was necessarily true that at least one of the four *target improvements* had to be included in the 1st – 4th rankings, since there were only 3 other improvements. Of the 136 participants who only ranked one of the *target improvements* 1st – 4th most important, 51 (approximately 8.5% of all participants) ranked the other improvements as 1st – 3rd most important, implying that all four *target improvements* were overall ranked less important for these participants. Only 1 participant ranked the four *target improvements* 1st – 4th most important.

Taken together, the findings from Stage 1 imply that the *target improvements* were on average ranked less important than the other improvements. Amongst the *target improvements*, External Sewer Flooding was ranked on average as most important. Since the later stages involve a novel mechanism designed to elicit the trade-offs between investment in different services, the results from this stage may be used as a benchmark for consistency in the later stages. Parametric analysis of the Ranking task, incorporating demographic effects, can be found in Appendix D.

It is also of interest to compare the results from this stage with the findings of the previous Scottish Water (2017) study, which asked citizens to rank similar service issues using a different methodology. Generally, there is reasonably high agreement between the results of the two studies. This study was not intended to be directly comparable to the 2017 study, and differences in the relative descriptions of the issues across the two studies may explain some disparities between the two sets of findings, as might the specific issues used.

For example, perhaps the most striking difference is for Interruptions to Supply across the two studies. However, there were clear distinctions between the specific issues present across the two tasks that make direct comparisons inappropriate. For example, the issues ranked worse than Short-Term Interruptions to Supply in the 2017 study were either not present in this present study or framed in such a way that might make them seem more important issues in this present study (e.g. the description of Beach/ River Cleanliness in this study refers explicitly to potential pollution issues, whereas in the 2017 study the issue was framed only in terms of the awarding of cleanliness ratings).

4. <u>Stage 2</u> – Basic Slider Task

Research Question: *What are citizens' general preferences for different specific service improvements?*

4.1. Introduction

In this stage, participants were asked to allocate an unknown level of investment between four of the improvements that they had just ranked in Stage 1. For each participant, the four improvements to be allocated investment in this stage were the same four *target improvements* from Stage 1. Participants were not provided a reason for why these improvements had been selected, and nor were they informed that all participants saw the same four improvements.

Participants were able to indicate their desired allocation of the unknown investment using a series of sliders, one for each of the four improvements. Figure 1 below provides a screenshot of the Basic Slider task environment, as seen by participants.

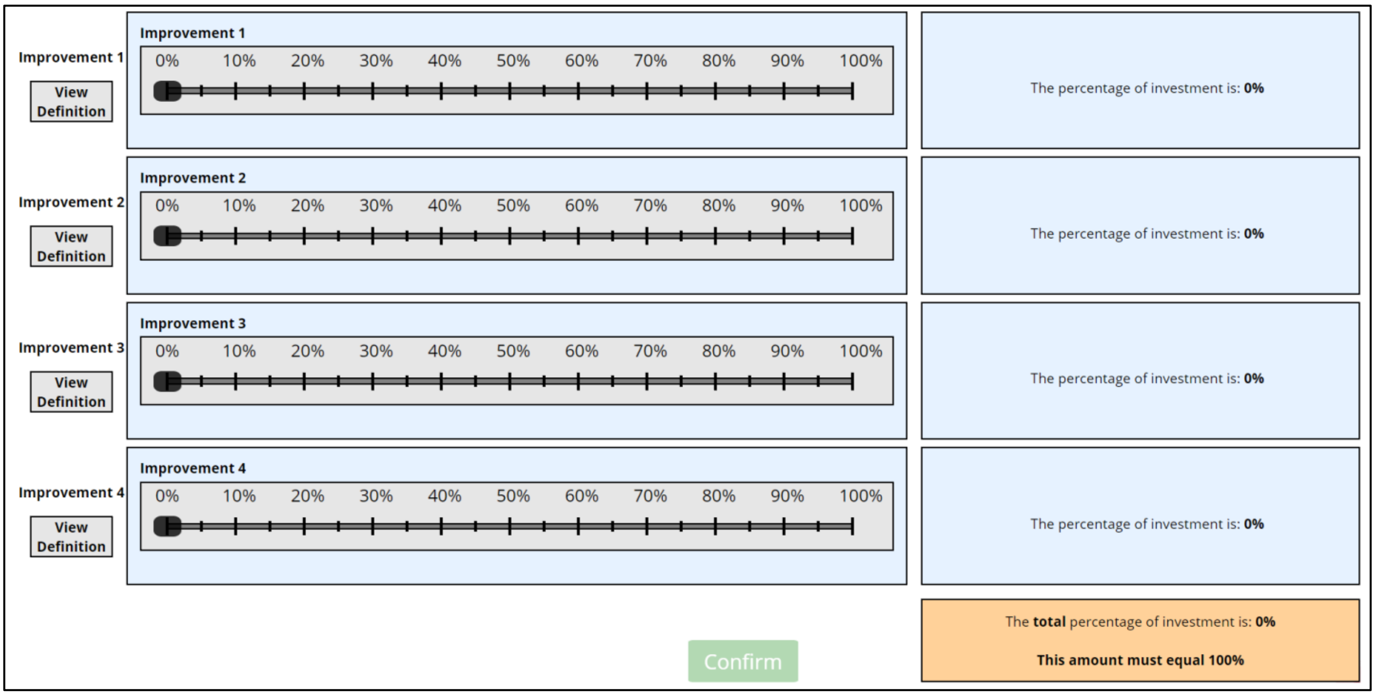


Figure 1. Example screenshot of the Basic Slider task environment

For each slider, participants could allocate 0% (which indicated that none of the investment would be spent on the improvement for that slider), 100% (which indicated that all of the investment would be spent on the improvement for that slider, and no investment would be spent on the other three improvements) or any percentage amount between 5% - 95%, in 5% increments. The total sum of the four slider allocations was required to equal 100%, and this total amount automatically updated on the participants' screens as they adjusted each slider.

The slider tools used in this study are a novel methodological design. Previous studies investigating citizen preferences for water services have commonly used a stated choice, or discrete choice, methodology (e.g. Hensher *et al*., 2005, Willis *et al*., 2005, Lanz and Provins, 2015). In these previous experiments, participants were faced with different scenarios comprising different water service combinations and costs, and asked to choose which scenarios they would prefer. Well-designed combinations of scenarios combined with assumptions about the shape of individual preferences enable a quantification of the estimated value of each service. However, as these combinations of scenarios are presented arbitrarily to participants, the trade-off between different attributes, and costs, may be implicit for participants as they respond to each decision. The design of these sliders was intended to make salient the trade-offs required in investment decisions. For a fixed level of investment, a certain amount of improvements may be feasible. In order to invest more in one improvement, the amount of investment available to be spent on another improvement must be reduced.

4.2. Results

If participants had equal preference for all of the four improvements, the allocation should be 25% for each improvement. In total, only 4.2% of participants indicated equal investment for all improvements. Only 0.5% of participants allocated 100% of investment to one improvement, and only 6.0% of participants allocated 0% to one (or more) improvement. The average difference between the highest and lowest allocation amount was 29.6 percentage points. In general, therefore, the setting of the sliders indicates that participants engaged positively with the task, providing a variety of investment allocations across all improvements. The order of presentation of improvements was randomised across participants, and there was no evidence of order influencing allocations in this task (details can be found in Appendix I).

Figure 2 below reports the average allocation by improvement. On average, 28.7% was allocated to investment in External Sewer Flooding, 27.1% to Renewable Energy, 22.5% to Interruptions to Supply and 21.8% to Rural Supplies.

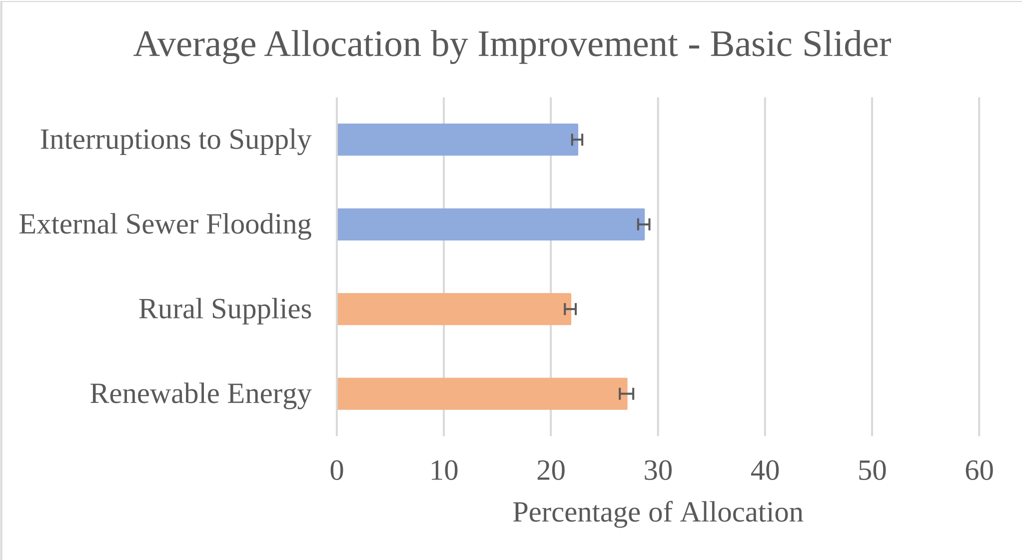


Figure 2. Average allocation by improvement in the Basic Slider task (error bars report standard errors) (blue colour coding denotes *personal* improvements, orange colour coding denotes *societal* improvements)

Figure 3, below, pools responses by the two broad improvement categories, *personal* and *societal*. Overall, there is a very slight preference on average for *personal* improvements (51.1%) over *societal* improvements (48.9%). 20.2% of participants allocated 50% to each group. Of those who gave an unequal allocation, there was no statistically significant difference in the likelihood of allocating more to *personal* improvements (52.9%) or *societal* improvements (47.1%) (χ^2= 1.640, p= 0.200)[2]. When comparing demographic effects on *personal*/ *societal* improvements, there is strong evidence of an age effect, with younger participants (i.e. those aged between 18-40) allocating significantly more to *societal* improvements than both participants aged 41-60 (p< 0.001) and those aged 61+ (p< 0.001). Additional details can be found in Appendix K. Whether this is reflective of younger citizens in general, suggesting that younger respondents will become less pro-social as they age, or whether this represents a cohort effect, where there has been a recent societal shift in attitudes, so that presently younger respondents will maintain their pro-social attitudes as they age, has potentially important implications from a policy perspective. In addition there were strong platform effects – with face-to-face respondents also allocating significantly more to *societal*

[2] There was also no significant difference between the allocation towards *societal* improvements among those participants who indicated that they are (45.6%) or are not (49.0%) connected to private water supplies (t= 1.038, p= 0.300).

improvements that online respondents (p= 0.011). Additional details can be found in Appendix H.

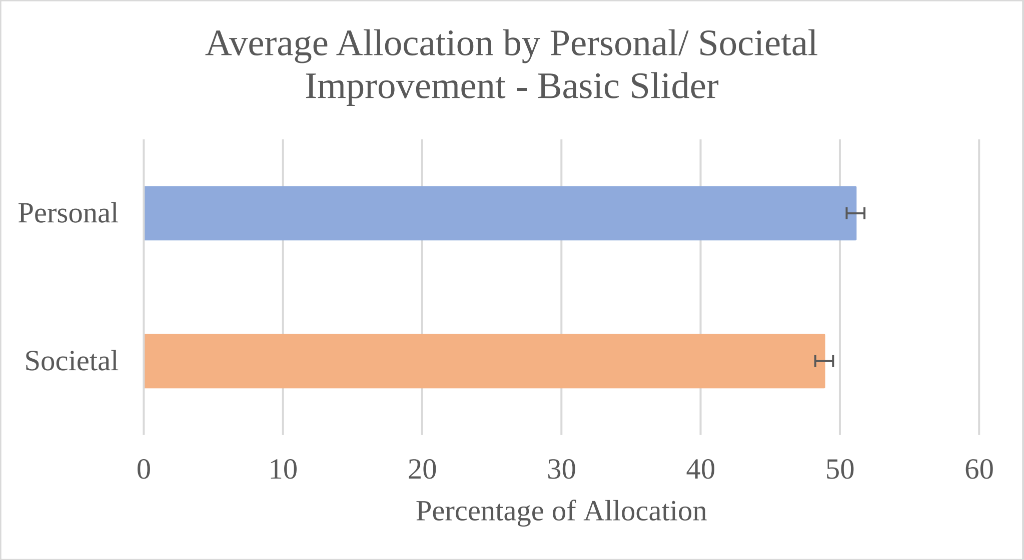


Figure 3. Average allocation by improvement groups the Basic Slider task (error bars report standard errors)

The use of a slider task was a novel methodology to elicit the trade-off decisions of consumers for a given set of investment priorities. As such it is useful to compare the decisions made in this task with the decisions made in the more conventional Ranking task. Evidence of consistency in the relative priorities given to individual improvements would be an indication of a level of engagement with the slider task that is comparable with more conventional preference elicitation mechanisms.

4.3. Basic Slider Task and Ranking Task Consistency

It is important to compare responses between the Ranking task and Basic Slider task as a measure of consistency between the two task types. Figure 4 below compares the average allocation amounts based on the relative rankings of the *target improvements* in the Ranking task. The *target improvement* that was ranked highest relative to the remaining three was allocated on average 35.9% of investment in the Basic Slider task. The second highest was allocated on average 26.4% and the third highest was allocated on average 21.6%. The *target improvement* that was ranked lowest relative to the remaining three was allocated on average 16.1% of investment in the Basic Slider task. Evidently, the higher a given *target improvement* was ranked relative to others, the higher the allocation it received in the Basic Slider task.

Participants revealed a remarkable degree of consistency between Ranking task decisions and Basic Slider task allocations. Since participants did not know that the four *target improvements* would be used in the later stages, such consistency would have been difficult to achieve through simple memory recall alone. This suggests that the decisions made in the Basic Slider task were reflective of genuine preferences, consistent with those elicited in the Ranking task.

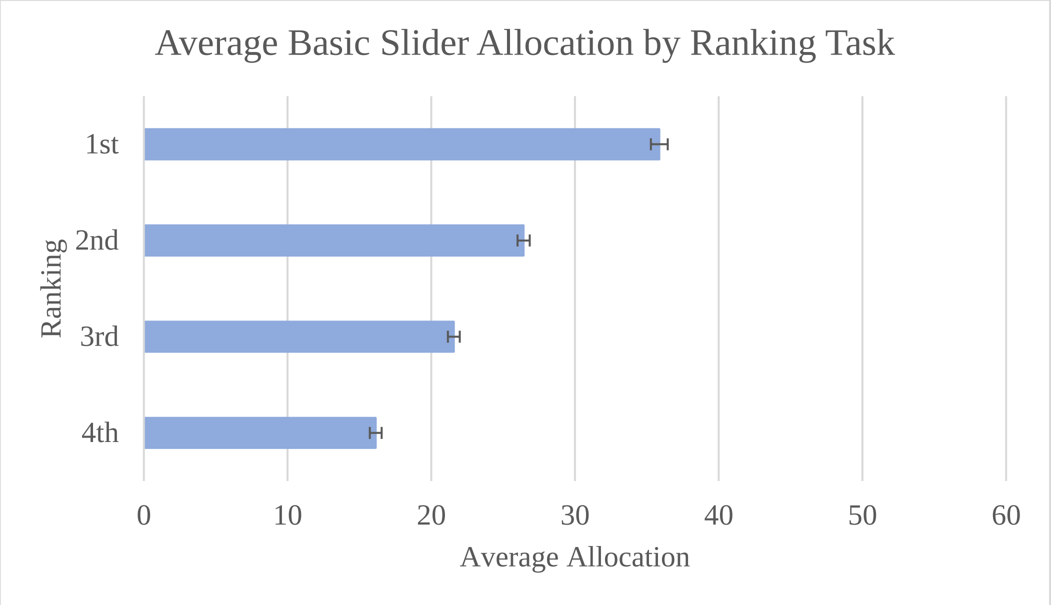


Figure 4. Average Basic Slider allocation by ranking in the Ranking task

5. <u>Stage 3</u> - Willingness-to-Pay Task

Research Question: *What would citizens deem to be an acceptable price rise for water charges to invest in specific service improvements?*

5.1. Introduction

In this stage participants were asked the maximum they would be willing to increase their current water bill by if they knew that the additional revenue raised would be used to improve the *target improvements* from the previous stage. There was no indication of the cost of improving each possible improvement in this stage. In order to generate more realistic responses, participants were first asked the council tax band of the property they lived in. This enabled the experiment to calculate each participants' actual annual water charge. Details of this can be found in Appendix J.

Participants were then returned to the basic slider screen, with an additional tool which enabled them to indicate the maximum amount they would be willing to add to their water charge. A screenshot of the tool is provided below in Figure 5. Participants could increase their water charge in £1 increments, or they could choose to add nothing to their water charge. Participants were informed that they were not able to reduce their water charge to less than its current level. The participants' new annual water charge (accounting for any increases) updated every time the participant used the tool, in nominal terms (i.e. not accounting for inflation). It was explained to participants that this represented a one-off price rise which would determine their new annual water charge. Previous research has shown that consumers may be averse to price rises, particularly if they perceive them to be unfair (Kahneman *et al*., 1986).

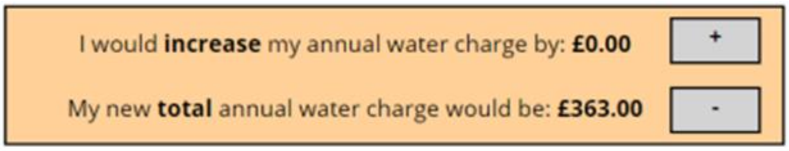


Figure 5. Screenshot of the willingness-to-pay tool used in Stage 3

The investment allocations participants had provided in the Basic Slider task were fixed in this stage, but every time the participant altered the willingness-to-pay tool, the monetary amount to be directed to each of the four improvements (in accordance with the allocation decisions of the Basic Slider task) were updated. For example, if a participant allocated 40% of investment to one improvement in the Basic Slider task, and indicated a willingness-to-pay an additional £10 on their water charge in this stage, then they were informed that £4.00 would be invested in that improvement.

It is important to acknowledge that the main aim of this stage was to provide a benchmark willingness-to-pay from which comparisons could be made when additional information (such as specific improvement costs and time horizons) was made available to the same group of individuals. This was not a controlled willingness-to-pay task aimed at explicitly eliciting the exact valuations of citizens. It is argued that open-ended questions such as the one used in this study may be less reliable than closed questions (i.e. "would you be willing to pay *X*?") (e.g. Arrow *et al*., 1993). However this would have made the dynamic design of later stages impractical, and since it was the relative changes in willingness-to-pay across stages that was a key outcome variable, this open-ended style question was deemed most appropriate.

5.2. Results

On average, participants indicated a willingness to pay an additional £11.03 (*SD*: 17.94) on their water charges, if this additional investment was to be spent in improving the four *target improvements*. Once accounting for current annual charges of each participant, this represents a willingness to increase water charges by 2.5% on average. However, 40.1% of participants indicated a willingness-to-pay of £0, suggesting that they would not be willing to add any additional amount to their water charge. Amongst those who were willing to pay something, average willingness-to-pay was £18.41 (*SD*: 20.03), a 4.2% increase on average.

As is common with findings of willingness-to-pay, especially for non-market goods, the responses were strongly right-skewed (i.e. lots of lower values and fewer very high values). 9.2% of responses indicated a willingness-to-pay of greater than £30, with one participant indicating a willingness-to-pay of £200. An analysis of the distributions of willingness-to-pay responses can be found in Table 3 below.

	All	
	(n)	**(%)**
£0	240	40.1%
£1 - £5	97	16.2%
£6 - £10	74	12.4%
£11 - £20	95	15.9%
£21 - £30	38	6.3%
£31 - £40	14	2.3%
£41 - £50	17	2.8%
£51 - £75	20	3.3%
£76 +	4	0.7%

Table 3. Distribution of willingness-to-pay responses

Although the very high responses are outliers relative to the overall population of responses, this required extensive use of the willingness-to-pay tool (since it only increased in £1 increments), and so these responses could not have been made through erroneous use of the tool. It is possible that responses in this task are a conservative estimate of participants' true preferences, if they became disinterested with engaging with the willingness-to-pay tool (where greater increases necessarily required greater effort). Additionally, framing each incremental price rise as a £1 increase might have signalled a lower expected price rise than if each unit increase was of a greater increment. In light of this it bears repeating that this stage aimed to provide an approximated baseline willingness-to-pay from which later willingness-to-pay decisions (in the presence of additional information) could be compared.

Parametric analyses of all Willingness-to-Pay tasks in this experiment can be found in Appendix G. Significant differences in willingness-to-pay were found across experimental platform. In face-to-face studies the amount that participants were willing to add to their water charge was over double that of those in online studies (additional details can be found in Appendix H).

It is possible that overall willingness-to-pay was influenced by a general attitude towards the *target improvements*. If the amount that citizens were willing to pay was influenced by how much they liked the improvements that were to be invested in, then this suggests that what participants are willing to pay for improvement depends on which improvements are to be invested in. Figure 6 below reports the willingness-to-pay of participants separated by how many *target improvements* were ranked within the top four of the Ranking task for each participant.

In general there appears evidence of an increasing trend of willingness-to-pay with increasing number of *target improvements* ranked within the top four of the Ranking task. Only one participant ranked the four *target improvements* as their top four most important in the Ranking task, so these were omitted from analysis (this participant provided a willingness-to-pay of £0). On average, those with one *target improvement* ranked in the top four were willing to pay £7.57, significantly less than those with two, (£11.74) (t= 2.605, p= 0.010), and significantly less than those with three (£13.20) (t= 2.194, p= 0.029). There was no significant difference between two and three (t= 0.705, p= 0.481).

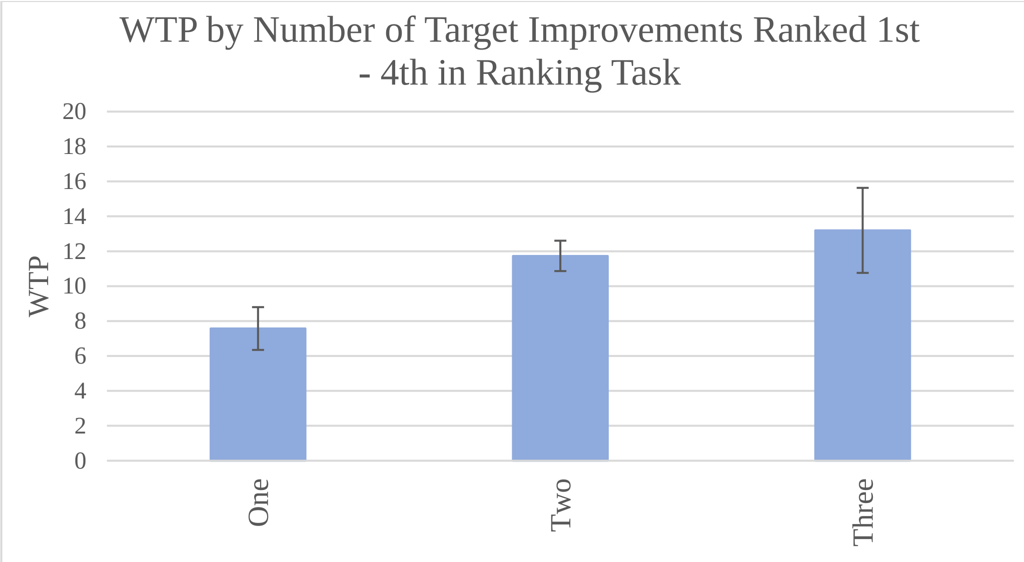


Figure 6. Willingness-to-pay by number of *target improvements* ranked 1st – 4th in the Ranking task

6. <u>Stage 4</u> – Cost Revelation Task

Research Question: *How does providing the costs of different specific service improvements alter the priorities and willingness-to-pay towards improvements of citizens?*

6.1. Introduction

In the previous stages, participants were not provided with any explicit information about the approximate costs of each improvement or the extent of the benefit that might be expected in return. This stage aimed to measure if, and how, participants would alter their preferences in the light of information about the different costs of different improvements. Preferences might change in two ways. First, knowing the absolute costs of improvements could influence the total willingness-to-pay of participants. Second, knowing the relative costs of improvements could influence the allocations provided to each improvement.

The task in this stage was a combination of both the Basic Slider task and the Willingness-to-Pay task. Participants were shown the same user interface as in these previous stages, but were also provided the unit cost of each improvement. The unit cost was presented as the cost per household to increase one unit of improvement *if every household in Scotland were to pay that additional cost*. Calculations of these costs were agreed in advance with stakeholders in the Scottish water industry and were broad approximations of the true costs of these improvements.

They were the same for all participants. The unit measure of improvement and its cost, for each improvement, are presented in Table 4 below.

Improvement	**Cost**	**Unit Measure**
Interruptions to Supply	2.5p	Reduce one incident of short term interruption to supply
External Sewer Flooding	10p	Reduce one external sewer flooding incident
Rural Supply	5p	Connect one rural home to public water system
Renewable Energy	50p	Generate one additional GWh of energy from renewable sources (where 1 GWh is equivalent to the annual electricity usage in 300 homes)

Table 4. Unit measure of improvement and corresponding cost for each improvement

In this stage, participants were asked to do two things:

1. Indicate the allocation of investment given to each improvement (as with the Basic Slider task)
2. Indicate their willingness-to-pay additional water charges (as with the Willingness-to-Pay task)

The sliders for each improvement were reset to 0% (therefore participants had to re-allocate the investment – there was no indication of the allocation they had provided in the Basic Slider task). The annual water charge was maintained at the level that was set in the Willingness-to-Pay task, but participants were able to increase or decrease this (but still could not decrease this to below their existing real annual water charge).

For each improvement, participants were also informed of how many units would be improved based on their total willingness-to-pay and their allocation to each investment. For example, if a participant allocated 40% to Renewable Energy improvements, and was willing to add £10 to their water charge (i.e. £4 to be spent on this improvement), they would be informed that this would allow an additional 8 GWh of energy generated by the Scottish water industry to be generated from renewable sources. These calculations updated on the participants' screens every time they adjusted a slider or the willingness-to-pay tool. Participants were encouraged to adjust these until they were happy with both their allocation and their willingness-to-pay.

6.2. Results

6.2.1. Slider Task

Figure 7 below reports the average allocations for each improvement in the Cost Revelation task. Comparing responses to allocations in the Cost Revelation task with the Basic Slider task reveals that 25.7% of participants provided the same allocations for all improvements in both tasks. For all improvements, between 84.6% - 88.3% of all allocations in this task were within 10 percentage points (either greater than or less than) of the allocation for the equivalent improvement in the Basic Slider task.

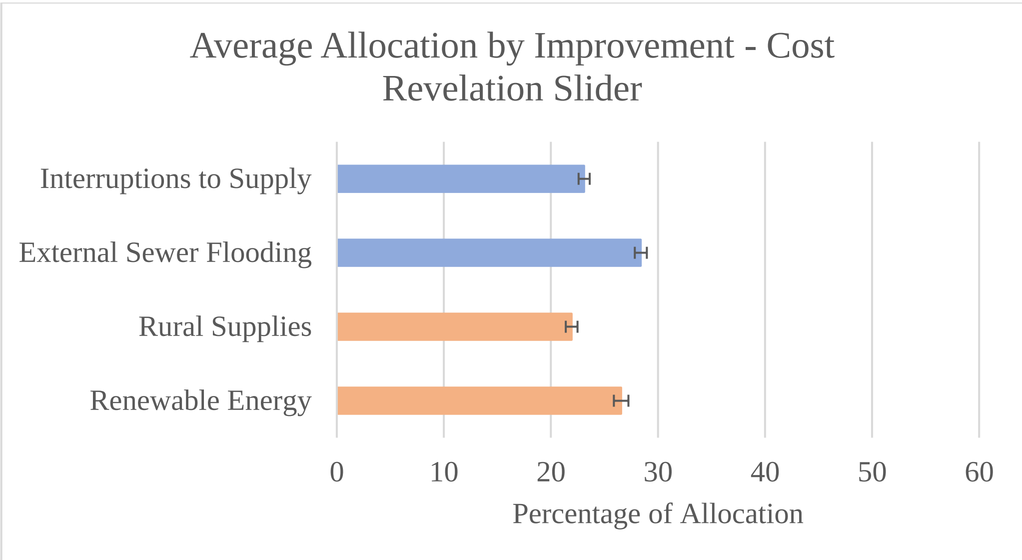


Figure 7. Average allocation by improvement in the Cost Revelation task (error bars report standard errors)

Figure 8 below reports the average difference in allocation for each improvement between the Basic Slider task and the Cost Revelation task. On average, the changes in allocation follow a pattern that would be expected if participants were readjusting allocations according to cost. However, the magnitude of these differences is small, and none of the differences in allocation amounts between tasks were statistically significant. The cheapest per unit improvement, Interruptions to Supply was allocated 0.7 percentage points more in the Cost Revelation task (t= 1.592, p= 0.112), and the second cheapest improvement, Rural Supplies was allocated 0.1 percentage points more in the Cost Revelation task (t= 0.282, p= 0.778). The third cheapest improvement, External Sewer Flooding was allocated 0.3 percentage points less in the Cost Revelation task (t= -0.569, p= 0.570), and the most expensive improvement, Renewable Energy, was allocated 0.5 percentage points less in the Cost Revelation task (t= -1.118, p= 0.264). However, the lack of statistically significant differences in allocations indicates that there is little evidence to suggest that participants' perceptions of investment prioritisation was strongly influenced by the relative costs of different improvements. Parametric analyses of the differences in allocation across the Basic Slider and Cost Revelation task may be found in Appendix E.

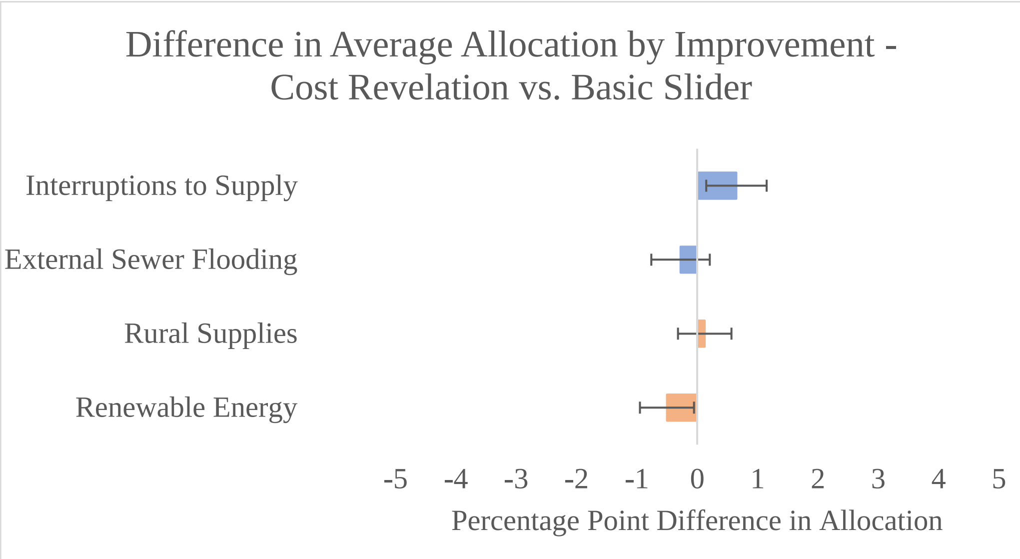


Figure 8. Difference in average allocation by improvement between Cost Revelation task and Basic Slider task

6.2.2. Willingness-to-Pay Task

In addition to adjusting the sliders, participants were also able to readjust their willingness-to-pay additional water charges in light of the cost revelation in this task. 84.1% of participants did not adjust their water charge between tasks. 1.2 % of participants reduced their water charge when the cost of improvements was revealed, and 14.7% of participants increased their water charge when the cost of improvements was revealed.

Overall, participants in this task were willing to pay an additional £12.09 (*SD*: 19.54), an increase of £1.06. The difference in willingness-to-pay between Stage 3 and Stage 4 was statistically significant (t= 4.787, p< 0.001). Thus, when participants were shown the cost of improvements, this significantly increased the additional water charge they were willing to tolerate, albeit that the effect was driven by a minority of individuals. Parametric analyses of all Willingness-to-Pay tasks in this experiment can be found in Appendix G.

This finding may have been connected to the way in which the costs of improvements were framed. By decomposing costs to the cost *per household* as opposed to the overall cost per unit increase, the additional cost per unit of improvement for each household appeared very small, and every additional £1 added to an individual's water charge was aggregated for all households in Scotland. This aimed to remove the temptation for any given participant to free-ride (i.e. refusing to tolerate a price increase in the hope that others would agree to an increase in their annual water charge). However, it is possible that framing the cost of improvements differently (for example in terms of absolute cost), might have generated different responses.

7. Stage 5 – Time Horizon Task

Research Question: *How does providing the time horizons of different specific service improvements alter the priorities and willingness-to-pay towards improvements of citizens?*

7.1. Introduction

In the previous stages, participants were not provided with any information about the length of time it may take for the benefits of each improvement to be realised. This stage aimed to measure if, and how, participants would alter their preferences in light of information about the differences in the time it would take for the benefits of different improvements to be felt.

The Time Horizon task was very similar to the Cost Revelation task. It was a combination of both the Basic Slider task and the Willingness-to-Pay task. In addition to the unit cost information provided in the Cost Revelation task, in this task participants were also informed of the estimated length of time it would take for the benefits of each improvement to be felt. Research into time discounting suggests that individuals give less weight to future events than immediate ones (Laibson, 1997). For example, one study found that individuals would accept a reward now that was four times smaller than a reward given in a years' time (Thaler, 1981).

In this stage the time horizons were manipulated in a between-subject design – that is, each participant saw only one version of time horizons, and this differed across participants. Intuitively, some of the specific improvements would feasibly take longer to improve than others. Because of this, the time horizons were not purely randomised across all improvements.

Two shorter time horizons were randomly assigned to Interruptions to Supply and External Sewer Flooding, and two longer time horizons were randomly assigned to Rural Supplies and Renewable Energy. The four time horizons were "Immediately", "5 Years", "10 Years" and "25 Years". These were determined in light of existing literature on time discounting of individuals. Although the time difference between "Immediately" and "5 Years", and "5 Years" and "10 Years" is identical, research suggests that attitudes across the two pairs of time horizons may not be equivalent. In addition, many improvements in the water industry may be inter-generational. A time horizon of "25 Years" aimed to elicit attitudes towards improvements which may be most benefitted by future generations of citizens. In total, there were four unique combinations of time horizons for improvements. These four options, one of which each participant was randomly assigned, are illustrated in Table 5, below.

Improvement	**Option 1**	**Option 2**	**Option 3**	**Option 4**
Interruptions to Supply	Immediately	5 Years	Immediately	5 Years
External Sewer Flooding	5 Years	Immediately	5 Years	Immediately
Rural Supplies	10 Years	10 Years	25 Years	25 Years
Renewable Energy	25 Years	25 Years	10 Years	10 Years

Table 5. List of all combinations of time horizon options

As in the Cost Revelation task, the sliders for each improvement were reset to 0% and the annual water charge was maintained at the level that was set in the Cost Revelation task, but participants were able to increase or decrease this (but still could not decrease this to below their existing real annual water charge).

7.2. Results

7.2.1. Slider Task

Since the time horizon of specific improvements differed across participants in this task, comparisons of allocations by time horizon must be made against the equivalent improvement in the Cost Revelation task for each participant. For example, if we were interested in the effect of the "5 Years" time horizon on responses, the allocation of all participants for whom Interruptions to Supply was given a "5 Years" time horizon would be compared with their allocation to Interruptions to Supply in the Cost Revelation task. Similarly the allocation of all participants for whom External Sewer Flooding was given a "5 Years" time horizon would be compared with their allocation to External Sewer Flooding in the Cost Revelation task. Figure 9 below outlines the average allocation in the Time Horizon task by different time horizons.

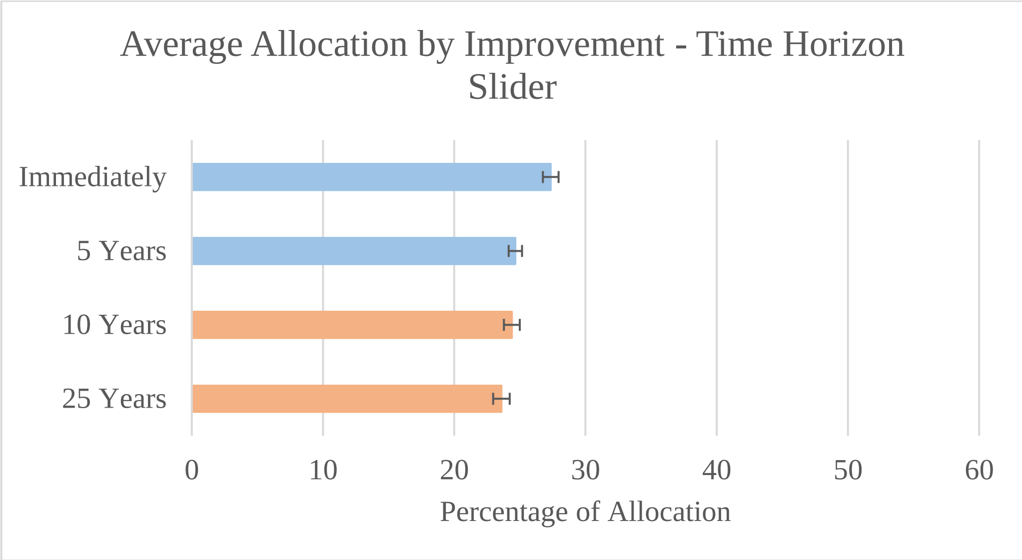


Figure 9. Average allocation by time horizon in the Time Horizon task (error bars report standard errors)

For an accurate comparison of the effect of revelation of time horizon, we compare the allocations to improvements in the Time Horizon task with the equivalent improvement in the Cost Revelation task for each participant. 31.6% of participants provided the same allocations for all improvements in both tasks. For all improvements, between 88.0% - 89.7% of all allocations in this task were within 10 percentage points (either greater than or less than) of the allocation for the equivalent improvement in the Cost Revelation task.

Figure 10 below reports the average difference in allocation for each equivalent improvement between the Cost Revelation task and the Time Horizon task. There is evidence of statistically significant allocation change when time horizons are revealed to participants. For improvements that were to be improved "Immediately", allocations increased by 1.4 percentage points in the Time Horizon task (t= 3.311, p= 0.001). For improvements that were to be improved in "5 Years", allocations decreased by 0.9 percentage points in the Time Horizon task (t= -2.412, p= 0.016). For improvements that were to be improved in "10 Years", allocations decreased by 0.1 percentage points in the Time Horizon task, but this decrease was not statistically significant (t= -0.263, p= 0.793). For improvements that were to be improved in "25 Years", allocations decreased by 0.4 percentage points in the Time Horizon task, but this decrease was not statistically significant (t= -1.137, p= 0.256). Parametric analyses of the differences in allocation across the Cost Revelation and Time Horizon task may be found in Appendix F.

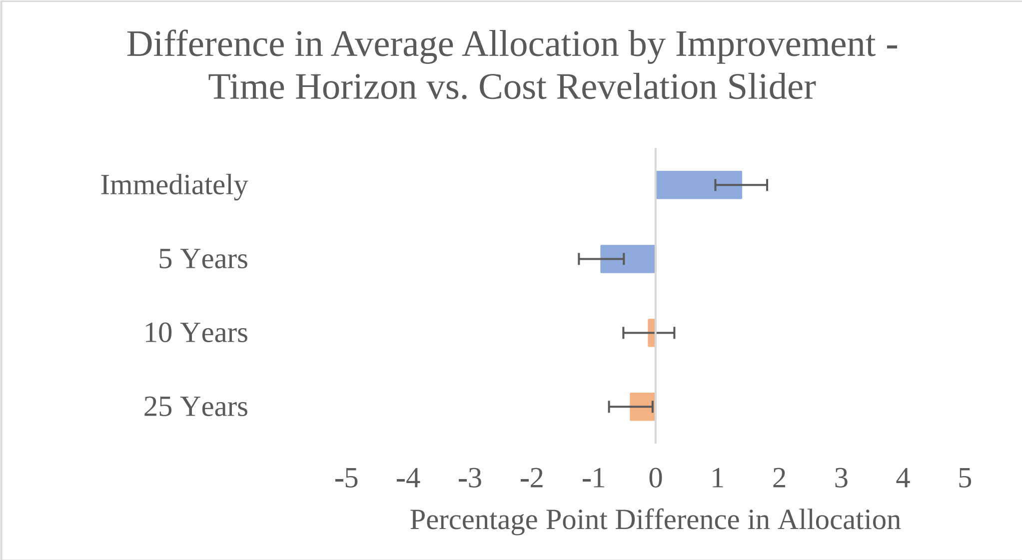


Figure 10. Difference in average allocation by equivalent improvement between Time Horizon task and Cost Revelation task

Existing research on time-discounting suggests that individuals may have a preference for more immediate rewards. The finding of increased allocation towards "Immediately" improvements supports this. What makes this slider design novel in the context of forcing decision trade-offs is that in order for participants to indicate a preference for "Immediately", this must be offset by a reduction in allocation elsewhere.

If time-discounting was a linear process, one might expect this to be disproportionately taken from the later time horizons instead. Theories of hyperbolic discounting suggest that discounting is not linear, but occurs more sharply for earlier delays and more slowly as the time delay increases. Depending on the slope of the hyperbolic discount function, the time durations of 5/10/25 years in the future may be expected to be more or less weighted by each individual. Whilst there was a significant increase in the allocation towards the "Immediately" time horizons, when comparing the average decreases across the three later time horizons, there were no statistically significant differences between "5 Years" and "10 Years" (t= -1.228, p= 0.220), "5 Years" and "25 Years" (t= -0.806, p= 0.421) or "10 Years" and "25 Years" (t= 0.484, p= 0.629). This suggests that citizens appear to disproportionately care about the more immediate future, but do not discern between medium and long-term improvements (i.e. 5/10/25 years in the future).

7.2.2. Willingness-to-Pay Task

As in the Cost Revelation task, participants were able to adjust both the sliders and their willingness-to-pay additional costs to their annual water charge in this task. 91.2% of participants did not change their willingness-to-pay between the Cost Revelation and Time Horizon tasks, 7.5% increased their willingness-to-pay once the time for improvements to be felt were revealed, and only 1.3% of participants reduced their willingness-to-pay with this information. Overall, participants in this task were willing to pay an additional £12.57 (*SD*: 19.99) on their annual water charge. This represents a £0.47 increase compared to the Cost Revelation task, a statistically significant increase ($t = 3.657$, $p < 0.001$). Parametric analyses of all Willingness-to-Pay tasks in this experiment can be found in Appendix G.

8. Summary of Findings

This experimental study aimed to better understand the preferences of Scottish citizens with regards to their priorities for future investment in the Scottish water industry. Through multiple stages we utilised a novel experimental methodology to elicit the attitudes of citizens regarding the importance of different service improvements, the degree to which future investment for specific service improvements should be prioritised, and how much should be spent to improve these services. We found evidence that the type of information provided influences attitudes to investment priorities and amounts. In this section we summarise the key results, and discuss potential underlying psychological mechanisms driving these responses.

The findings of this study can be summarised as five key results:

1. **There is substantial variation in the preferences of participants.** Whilst there is evidence of some systematic preferences across stages, at the individual level a variety of preferences for water improvements are reported. No one (or multiple) service improvements were consistently and substantially preferred to others in a systematic way across all participants across stages. However, at the individual level, these responses were consistent across stages (in particular across the Ranking task and Basic Slider task), suggesting that these preferences were not borne of random decision-making or disinterest of participants; rather participants appeared to be engaged in the

study but simply assigned a broad array of preferences towards different water improvements.

2. **Participants in general are willing to pay additional water charges to improve specific services.** Almost 60% of participants were willing to pay extra on their current annual water charges to improve the *target improvements*, without any information about specific costs or time horizons. Across all participants this amounted to £11.03 extra, a 2.5% increase in their current annual water charges over the next 12 months. This is a higher proportion of participants and a larger increase than recorded by a somewhat different method in Stage 1 of Belton *et al*. (2020). In that study, we elicited the short run priors of what constituted a reasonable price change over the next 12 month period, with no indication of how any additional funds would be spent. Unlike in the present study, participants could opt for a price fall. Excluding those who did (for comparability), the average acceptable price rise in Stage 1 of Belton *et al*. (2020) was calculated as approximately £2.47 (or £5.27 excluding also those who selected "No change")[3]. The disparity between the two studies, both in proportion opting for an increase and size of increase, suggests that providing information about how and where additional investment is to be spent may increase willingness-to-pay. Some caution is warranted, as the different responses could in part be due to the somewhat different questions asked. However, combining the results of both studies indicates a potential willingness to consider modest increases in charges, perhaps especially among those with better information. This difference between the two measures is also consistent with the more general behavioural tendency to prefer options that are concrete or familiar over options that are uncertain or unknown. On the assumption that learning about the possible benefits improved participant understanding, it can also be argued that the present results are more likely to reflect true underlying preferences (Beshears *et al*., 2008).
3. **Additional information about specific improvement costs and time horizons led to increased willingness-to-pay additional water charges.** Willingness-to-pay increased by approximately £1.06 when information about the specific costs of service improvements was provided. However this was driven by only 14.7% of participants

[3] Stage 1 of Belton *et al*. (2020) was a multiple choice question of *"Which of the following options do you believe to be the most acceptable price change for water charges over the next year?"*. Participants were able to select a response option between the ranges of "Decrease by £5.01 - £10.00" to "Increase by £15.01 - £20.00".

who increased their willingness-to-pay between tasks. The large majority, 84.1%, did not change their willingness-to-pay when provided with specific cost information and 1.2% reduced it. This suggests that, when citizens are told how much improvement can be made from their increased water charges, they are on average willing to pay more towards these improvements – although the effect is not large. The costs may appear cheaper than some citizens would have anticipated. The cost information in this study was presented at a per-household level (i.e. the cost to increase one unit of improvement *if every household in Scotland were to pay that additional cost*). This meant, for any additional £1 indicated by the participant, multiple improvements could be made. The framing of costs in this way was a fundamental aspect of the slider task design. It is possible that if these costs were framed differently, for example as larger, absolute unit costs, then participants' willingness to increase their water charges may have differed. Willingness-to-pay also increased again but to a lesser extent (£0.47) when participants were given information about the length of time it would take for improvements to be felt. It is possible that the time horizons provided were shorter than some participants might have anticipated, although they ranged from zero to 25 years away and were not based on specific estimates from the Scottish water industry. These findings may also be related to the general behavioural tendency to prefer things that are more familiar or concrete as suggested above – where increased learning about the costs and time horizons of improvements led to a preference (i.e. higher willingness-to-pay) when compared to willingness-to-pay without this information.

4. **Citizens have a preference for more immediate improvements.** In the Time Horizon task, participants allocated significantly more to the improvement whose benefits could be felt "Immediately". In isolation, this finding is perhaps to be expected. What was more surprising, however, was that we recorded no substantial differentiation in preferences between time periods of "5 Years", "10 Years" and "25 Years" in the future. This suggests that citizens perceive more immediate time horizons to be more important, but do not distinguish between increasing future time horizons. In general, however, the differences between the time horizons were quite small. It is possible that the design underestimated the effect of time horizons. The behavioural phenomenon of *anchoring* (e.g. Tversky and Kahneman, 1974) may have contributed if people anchored their allocations in later slider tasks to their earlier allocation decisions or, perhaps, wanted to appear consistent in their decisions. We attempted to reduce this by resetting every slider to 0% in each slider task, but it is reasonable to assume that participants could

remember (with some potential scope for error) their allocations across the sliders in the previous task. Nevertheless, we recorded only a relatively weak preference for earlier realisation of benefits and it may be that citizens are willing to take a long-term view of returns to investment in a resource like water.

5. **The experimental platform used can influence responses.** A more general finding of this study was of differences in the responses of online and face-to-face participants. This was observed in both the slider tasks and the willingness-to-pay task. In the slider tasks, face-to-face participants indicated a preference to allocate more to *societal* improvements, whereas online respondents indicated a preference to allocate more to *personal* improvements. In the willingness-to-pay task, face-to-face respondents were willing to increase their annual water charge by approximately double that of online respondents. It is possible that the types of people who are incentivised to attend face-to-face sessions are systematically different to those who complete online studies (even after controlling for observable demographic differences), but this is conjecture. Nevertheless, these findings support those of Belton *et al*. (2020) which indicate that running the same experiment on different platforms can lead to systematic differences in responses by platform type.

9. Policy Implications

This final section considers the implications of this second study in the experimental research series, designed to provide behavioural evidence to inform regulation of the Scottish water industry. The results build on and extend those obtained in Belton *et al*. (2020).

Although the evidence base concerning household preferences for water services is growing, it is important to bear in mind that much remains to be investigated. In particular, while the present study elicited preferences for different improvements after participants were informed of the costs and benefits involved, it did so without communicating information on baseline levels of service and associated risks. For instance, a relative preference for investment designed to reduce interruptions to supply might be altered by knowing the probability of such an interruption; the preference for cleaner beach and river water may be changed by understanding current levels of pollution, and so on. Thus, inferences made from the present findings need to be drawn cautiously, in the knowledge that additional relevant empirical evidence is likely to emerge as this research programme progresses and other results come to light. Nevertheless, the results have implications that warrant consideration.

It is important when considering this study to see the wood as well as the trees. The pattern of individual preferences for specific improvements varied greatly and participants generally allocated substantially more funds to their favoured improvement. Yet although statistically significant differences were recorded, the average effects – aggregated preferences – across the seven improvements (and indeed four *target improvements*) were not very strong. The relatively modest gap in average ranking (out of seven) between the highest and lowest priority improvement (3.24 versus 4.79) in the Ranking task and the fact that the aggregate investment allocations to the four target improvements lay within a narrow band in the Basic Slider task (21.8-28.7%), indicates a lack of consistent and strong aggregate preferences, at least for the improvements tested here based on the descriptions provided. It is important to recall that this was not due to lack of engagement with the task or a bias towards equal allocation, given that the majority of participants in the allocation task opted for a highly unequal allocation, suggesting a strong individual preference for some improvements over others. Rather, because different respondents had different priorities the outcome was a relatively even aggregate allocation. Note that this might not have been the case. A strong aggregate preference could easily have emerged from the study design, at any of the stages as new information was communicated. An aggregate preference could have clearly favoured one particular improvement or one type (personal, societal, environmental, etc.). That this did not occur is telling. Moreover, throughout the study we found no consistent preference for improvements that generated individual benefits over those that generated societal benefits, at the population level. From a policy perspective, this plurality and balance of views across different households is important to recognise. This is a straightforward and potentially important finding in the context of a regulatory system that, in part, aims to respond to the preferences of Scottish households.

That said, improvements to beach/river cleanliness and reductions in supply leakages were the biggest priorities of this representative sample of Scottish citizens. It is interesting that both of these improvements have an environmental dimension, although other environmental benefits were not ranked so highly. Also notable is the fact that the four highest ranked improvements all featured some visible component (e.g. water/sewage flooding, visible pollution or change in water appearance). It is possible the findings reflect a preference for improvements that can be observed.

Neither participants' individual responses nor the relatively even aggregate allocation of investment were strongly disrupted by providing information about the costs and benefits of the four target improvements. This result might be interpreted as an indication that the scales of the costs and benefits came as no surprise to participants. Alternatively, participants may have viewed their relative allocation as more than a merely transactional process of investing money for a return, but also as one that gave a signal about priorities. Either way, the relatively even preferences recorded across the four target improvements were robust to providing information about costs and benefits.

Making the costs and benefits of different improvements explicit did generate an aggregate increase in willingness to pay, via increased water charges. This was also not a foregone conclusion. Participants could instead have been disappointed when they discovered what they would get in return for their money and expressed this disappointment through a stronger preference for a zero increase. The response required an active decision to express a preference for an increase in charges and 60% of participants opted to do so (the remaining 40% indicated a desire for no increase, as there was no option to signal a desire for a decrease in charges). Broadly, the findings support those of Stage 1 in Belton *et al*. (2020), obtained via a different method, although willingness-to-pay in Study 2 was estimated to be somewhat greater. Thus, while the evidence records variation, both studies suggests that at the aggregate level there is some willingness among Scottish households to accept modest increases in charges. Explicit information in relation to what can be expected in return does not undermine this willingness and may in fact increase it.

The present study also varied the stated timescale over which benefits would come to fruition. Unsurprisingly, given existing evidence on time discounting, people expressed a preference for immediate benefits over delayed ones. While important to note, this effect was relatively small and accompanied by the perhaps more surprising finding that there was almost no difference in preference between benefits that would be felt over 5, 10 or 25 years. Arguably, this result is consistent with that relating to *personal* and *societal* improvements – if the improvement is mainly to be felt by others, how much does it matter whether it is others now or others later? The scale of time discounting we recorded was less than the steep discounting typically seen in behavioural studies in which individuals trade off immediate against future outcomes. It is possible that this outcome was related to the innovative slider method employed, which made the trade-off involved unavoidable. But it is also possible that the shallow time discounting recorded reflects an underlying recognition that the nation's water system needs to be reliable and sustainable across decades and generations. Thus, this finding is suggestive and might reflect a potential willingness to entertain investment in water services over long time horizons, but further research is required before such a conclusion might be regarded as solid.

Finally, the present study found some systematic differences in preferences between young and old people, most notably in respect of the desire for priority to be given to societal improvements. This raises an issue in relation to the faithful representation of citizens' preferences in decisions about the future of the industry. It is impossible to determine whether the difference is an effect of age or cohort, i.e., whether these pro-social preferences will weaken with age or whether the current generation of young people possesses stronger pro-social preferences that will persist. To the extent that the results reflect cohort rather than age, they raise an issue of intergenerational fairness. It is reasonable to ask whether equal weight should be accorded to the differing preferences of older and younger people when the latter will, ultimately, be more affected by long-term investment decisions taken at the present time.

10. References

Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., and Schuman, H. (1993), Report of the NOAA panel on contingent valuation, *Federal Register*, *58*

Belton, C. A., Robertson, D. A., and Lunn, P. D. (2020), An experimental study of attitudes to changing water charges in Scotland, *ESRI Working Paper Series No. 654*

Beshears, J., Choi, J.J., Laibson, D., and Madrian, B.C. (2008), How are preferences revealed?, *Journal of Pubic Economics, 92*, 1787-1794

Hensher, D., Shore, N., and Train, K. (2005), Households' willingness to pay for water service attributes, *Environmental and Resource Economics, 32*, 509-531

Irwin, J. R., Slovic, P., Lichtenstein, S., and McClelland, G. H. (1993), Preference reversals and the measurement of environmental values, *Journal of Risk and Uncertainty*, *6*, 5-18

Kahneman, D., Knetsch, J. L., and Thaler, R. (1986), Fairness as a constraint on profit seeking: Entitlements in the market, *American Economic Review*, 728-741

Laibson, D. (1997), Golden eggs and hyperbolic discounting, *Quarterly Journal of Economics*, *112*, 443-477

Lanz, B., and Provins, A. (2015), Using discrete choice experiments to regulate the provision of water services: Do status quo choices reflect preferences?, *Journal of Regulatory Economics*, *47*, 300-324

Scottish Water (2017), Research on customer priorities to inform our strategic review (2021), *Engage21: Customer Research*, 1-22

Thaler, R. (1981), Some empirical evidence on dynamic inconsistency, *Economics Letters*, *8*, 201-207

Tversky, A. and Kahneman, D. (1974), Judgment under uncertainty: Heuristics and biases, *Science*, *185*, 1124-1131

Water Industry Commission for Scotland (2017), Innovation, *Strategic Review of Charges 2021-2027 Initial Decision Paper*, *11*, 1-4

Walker, G. (2017), Untapped potential: Consumer views on water policy, *Consumer Futures Unit Publication Series*, *13*, 1-22

Willis, K., Scarpa, R., and Acutt, M. (2005), Assessing water company customer preferences and willingness to pay for service improvements: A stated choice analysis, *Water Resource Research*, *41*, 1-11

Appendices

Appendix A – Description of Service Improvements

Below is a list of the descriptions of service issues and corresponding potential improvements that participants were shown prior to the Ranking task.

Target Improvements

Short term interruptions to supply
- Water supply to homes is occasionally interrupted unexpectedly. This means that a property would be without water for a period of time.
- A short term interruption is classed as between 6-12 hours duration.
- This can affect daily routine in various ways e.g. the toilet cannot be flushed, there is no water in the tap, the shower cannot be used, for the duration of the 6-12 hour period.

Possible Improvement **- Reducing the number of short term interruptions to supply**

External sewer flooding incidents
- On occasion sewers do not have capacity to cope with the amount of water they need to carry and sewage can spill out of the system, leading to external sewer flooding.
- External sewer flooding means not being able to access your property, or any land on your property, without stepping through sewage flooding.
- This can include detached garages and outhouses, and also roads near your home or community areas.

Possible Improvement **- Reducing the number of external sewer flooding incidents**

Rural supplies to the water system
- A number of homes in remote areas currently have no access to the public Scottish water supply.
- These homes have to rely on their own private supply.
- Private supplies may be less reliable and of lower quality than publicly supplied water, and this may sometimes represent a health risk.

Possible Improvement **- Connecting rural supplies to the water system**

Non-renewable energy usage
- Powering the Scottish water supply requires substantial energy usage.
- Energy production from fossil fuels is a source of greenhouse gas emissions that contribute to climate change.
- Generating this energy from renewable sources instead can reduce the amount of greenhouse gases generated by the Scottish water supply.

Possible Improvement **- Increasing share of renewable energy**

Other Improvements

Leakage in supply pipes
- On occasion leakage in supply pipes can lead to water that is supposed to be supplied to properties being wasted.
- High levels of leakage may reduce the reliability of water supply.
- It may also result in visible water leakages rising to the surface in affected areas.

Possible Improvement **- Reducing leakage in supply pipes**

Drinking water look, taste and smell
- In older water mains pipes, some natural chemicals in water may leave deposits, leading to changes in colour or sediment in water in the home.
- Some chemicals used to keep water clean may interact with other natural chemicals in water and cause a different taste or smell to water.
- These effects are not harmful, but may make water look, taste, or smell different.

Possible Improvement **- Improving drinking water look, taste and smell**

Cleanliness of rivers, seas and beaches
- Many sewer networks are designed to spill some sewerage into rivers and seas.
- When operating properly these occasional discharges cause little risk to the environment.
- If a sewer capacity is too small, or the network is not operating properly, these discharges can impact water environmental quality or lead to visible pollution in rivers, seas and on beaches.

Possible Improvement **- Improving cleanliness of rivers, seas and beaches**

Appendix B – Wilcoxon sign-rank tests for pairwise comparisons of improvements

	Supply Pipe Leakages	**External Sewer Flooding**	**Drinking Water Quality**	**Renewable Energy**	**Rural Supplies**	**Interruptions to Supply**
Beach/ River Cleanliness	52.3% *z*= 1.156 *p*= 0.248	56.8% *z*= 4.322 *p*< 0.001	54.6% *z*= 4.291 *p*< 0.001	68.8% *z*= 10.208 *p*< 0.001	72.5% *z*= 10.718 *p*< 0.001	70.8% *z*= 11.401 *p*< 0.001
	Supply Pipe Leakages	56.6% *z*= 3.500 *p*< 0.001	54.1% *z*= 3.279 *p*= 0.001	63.4% *z*= 7.758 *p*< 0.001	68.1% *z*= 9.989 *p*< 0.001	72.5% *z*= 11.240 *p*< 0.001
		External Sewer Flooding	50.6% *z*= 0.457 *p*= 0.648	59.9% *z*= 5.225 *p*< 0.001	61.1% *z*= 7.819 *p*<0.001	68.1% *z*= 8.782 *p*< 0.001
			Drinking Water Quality	57.9% *z*= 4.324 *p*< 0.001	59.4% *z*= 6.285 *p*< 0.001	61.9% *z*= 8.066 *p*< 0.001
				Renewable Energy	55.6% *z*= 1.974 *p*= 0.048	53.9% *z*= 2.892 *p*= 0.004
					Rural Supplies	52.3% *z*= 1.440 *p*= 0.150

Table B.1. Wilcoxon sign-rank tests for pairwise comparisons of all improvements (reporting the percentage of participants who gave a lower ranking to the *row* improvement to each *column* pairwise comparison)

	Renewable Energy	**Rural Supplies**	**Interruptions to Supply**
External Sewer Flooding	59.9% *z*= 5.314 *p*< 0.001	61.1% *z*= 7.566 *p*< 0.001	68.1% *z*= 8.511 *p*< 0.001
	Renewable Energy	55.6% *z*= 1.622 *p*= 0.105	53.9% *z*= 3.102 *p*= 0.002
		Rural Supplies	52.3% *z*= 1.481 *p*= 0.139

Table B.2. Wilcoxon sign-rank tests for pairwise comparisons of *target improvements* (reporting the percentage of participants who gave a lower ranking to the *row* improvement to each *column* pairwise comparison)

Appendix C – **Analysis of the distributions of rankings of improvements in the Ranking task**

Histograms - Target Improvements - All

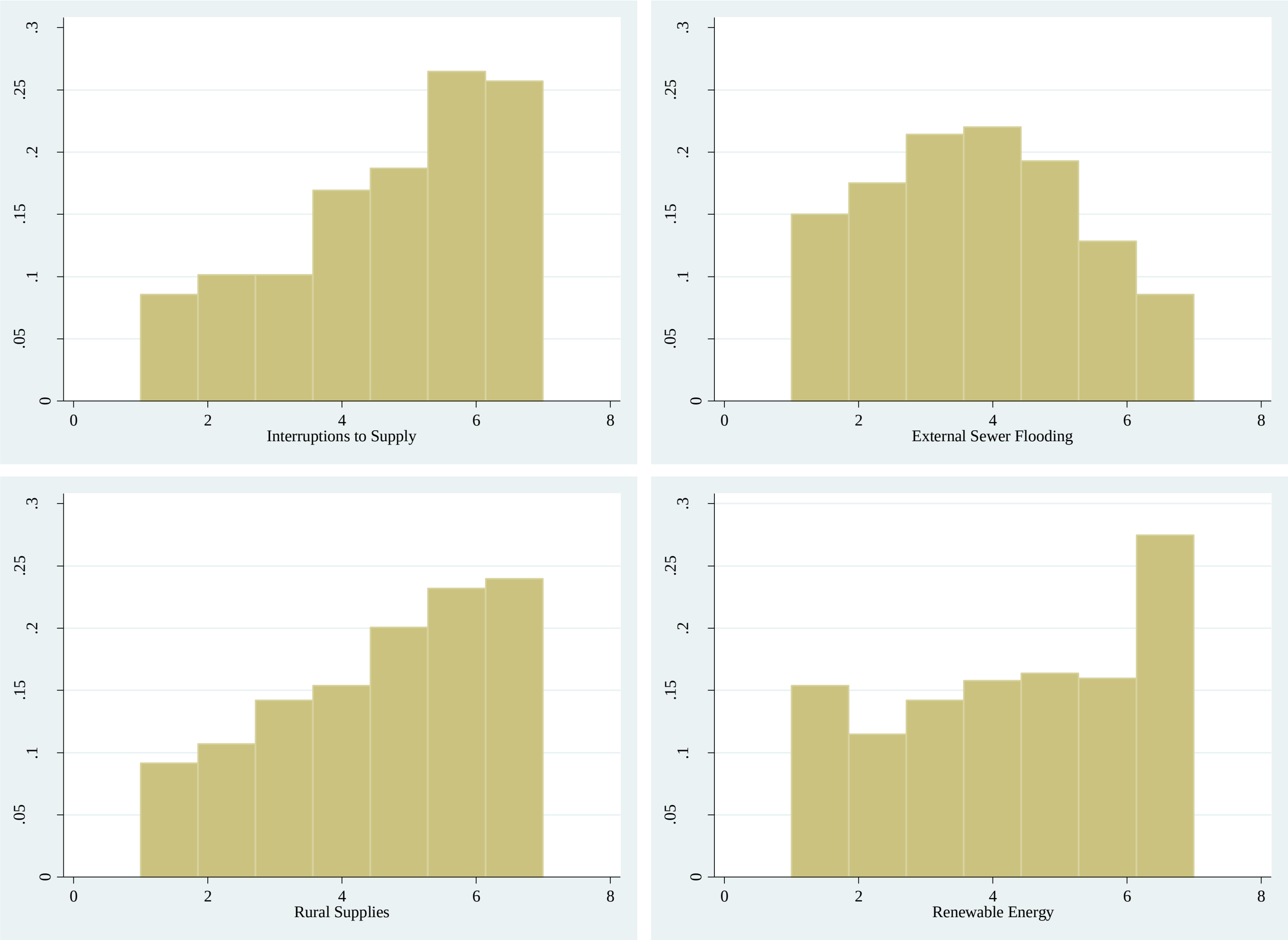


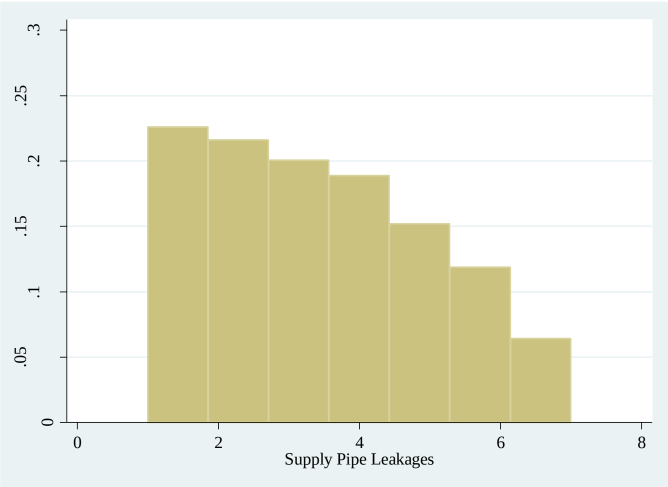
Figure C.1. Histogram of the rankings of Interruptions to Supply in the Ranking task (*top left*)

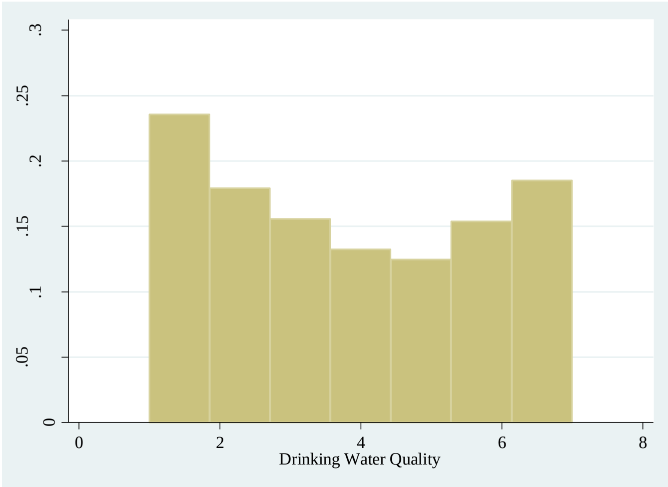
Figure C.2. Histogram of the rankings of External Sewer Flooding in the Ranking task (*top right*)

Figure C.3. Histogram of the rankings of Rural Supplies in the Ranking task (*bottom left*)

Figure C.4. Histogram of the rankings of Renewable Energy in the Ranking task (*bottom right*)

Histograms - Other Improvements – All





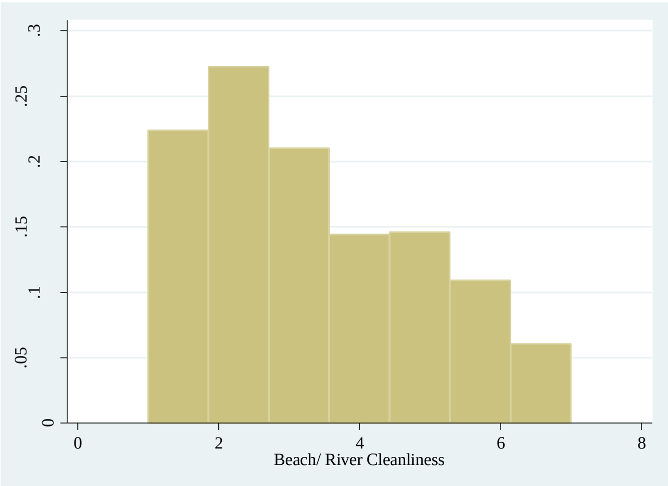


Figure C.5. Histogram of the rankings of Supply Pipe Leakages in the Ranking task (*top left*)

Figure C.6. Histogram of rankings of Drinking Water Quality in the Ranking task (*top right*)

Figure C.7. Histogram of rankings of Beach/ River Cleanliness in the Ranking task (*bottom left*)

Histograms - Target Improvements – Only

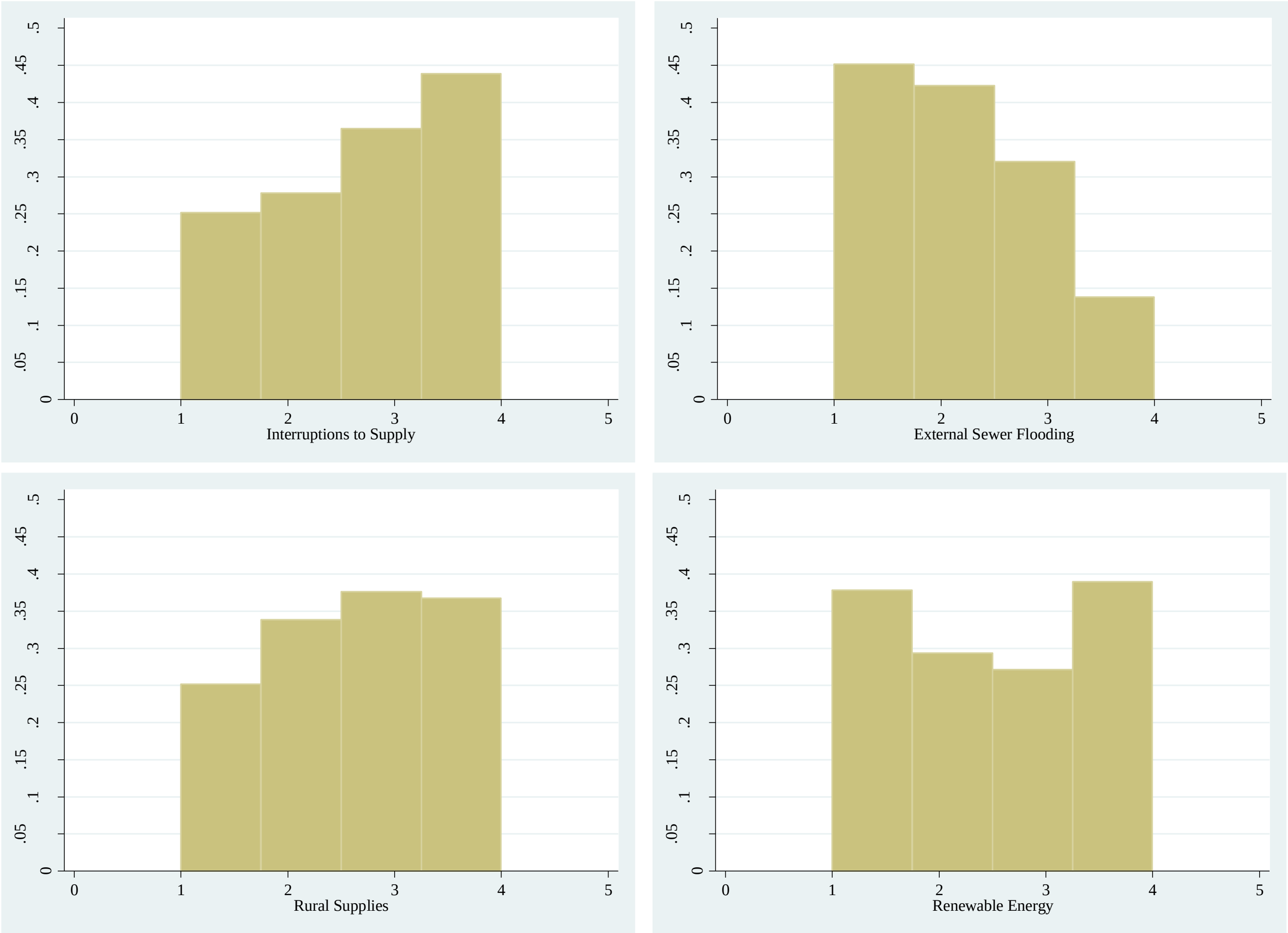


Figure C.8. Histogram of the rankings of Interruptions to Supply among *target improvements* only in the Ranking task (*top left*)

Figure C.9. Histogram of the rankings of External Sewer Flooding among *target improvements* only in the Ranking task (*top right*)

Figure C.10. Histogram of the rankings of Rural Supplies among *target improvements* only in the Ranking task (*bottom left*)

Figure C.11. Histogram of the rankings of Renewable Energy among *target improvements* only in the Ranking task (*bottom right*)

Table C.1 below reports the results of variance ratio tests for a number of pairwise comparisons of improvements. Both Renewable Energy and Drinking Water Quality indicate evidence of a U-shaped distribution. Results of variance ratio test for pairwise comparisons of improvements with Renewable Energy and Drinking Water Quality separately are provided below. To account for potential increased likelihood of Type I error (i.e. incorrectly rejecting a null hypothesis) which is associated with conducting repeated analyses, we test each individual hypothesis using a Bonferroni correction. The critical value is therefore 0.1/5 = 0.02, and null hypotheses are rejected for tests that report *p*-values below this critical value. As shown below, results suggest that variance is larger and statistically significant for these two distributions versus other improvements, with the exception of the comparison between Renewable Energy and Rural Supplies which falls just short of statistical significance. Overall this is indicative of a more varied response range across individuals for both Renewable Energy and Drinking Water Quality than other improvements.

Renewable Energy (*SD: 2.086*)	***F*-stat**	***p*-value**
Interruptions to Supply (*SD:1.876*)	1.236	0.010
External Sewer Flooding (*SD: 1.773*)	1.384	< 0.001
Rural Supplies (*SD: 1.898*)	1.209	0.021
Supply Pipe Leakages (*SD: 1.818*)	1.317	< 0.001
Beach/ River Cleanliness (*SD: 1.800*)	1.343	< 0.001
Drinking Water Quality (*SD: 2.144*)		
Interruptions to Supply (*SD:1.877*)	1.305	0.001
External Sewer Flooding (*SD: 1.773*)	1.461	< 0.001
Rural Supplies (*SD: 1.898*)	1.276	0.003
Supply Pipe Leakages (*SD: 1.818*)	1.390	< 0.001
Beach/ River Cleanliness (*SD: 1.800*)	1.418	< 0.001

Table C.1. Results of variance ratio tests for pairwise comparisons of improvements with Renewable Energy and Drinking Water Quality separately

<u>Appendix D</u> – Ranking Task Parametric Results

Here we report the results of statistical analyses which aim to isolate the individual effects of presentation, experience and demographic effects on participants' rankings for each of the seven possible service improvements in the ranking task. For each possible improvement we ran a separate ordered logistic regression. The dependent variable was the ranking (from 1 - 7) for that improvement, and presentation order, platform, experience and demographic characteristics were used as the independent variables, with log odds reported.

Table D.1 reports the results for the four *target improvements* and Table D.2 reports the results for the remaining three improvements. For ease of interpretation, Tables D.1 and D.2 do not report the coefficients for "I don't know"/ "Prefer not to say" responses to any experience or demographic questions, although these were controlled for in the models. The inclusion of multiple demographic variables leads to some models failing standard assumption checks for ordered logistic regression models. However, the use of more appropriate models do not substantively change outcomes and so ordered logistic regressions are used for ease of interpretation[4][5]. Any meaningful effects of presentation, platform, demographic or experience effects in this stage are reported in Appendices H - L.

[4] A Brant test for the proportional odds assumption could not be computed for all of the models as a result of the inclusion of "Prefer Not to Say" choices for demographic variables, where there was insufficient sample size to compute binary logits across all variables. Removing these from analyses enabled a Brant test to be computed for all but Model 2. There were insufficient observations at the extreme rankings for Model 2. Pooling rankings 1 and 2, and rankings 6 and 7 in Model 2 overcame this issue without substantively changing results.

[5] Models 1 (p= 0.268), 4 (p= 0.360), 5 (p= 0.281) and 7 (p=0.397) passed their corresponding Brant tests. Whilst all included variables in Models 2 (p= 0.015), 3 (p= 0.003) and 6 (p= 0.034) did not pass the proportional odds assumption, the sample size is sufficiently large and the difference in changes between ranking values does not influence interpretation of the key results, and so ordered logistic regression is a preferred model to a generalised ordered logistic regression (Williams, 2016).

Ranking Task	**Model 1** Interruptions to Supply	**Model 2** External Sewer Flooding	**Model 3** Rural Supplies	**Model 4** Renewable Energy
Order – First Page (*Ref: No*)				
Yes	-0.5575***	-0.7322***	-0.5736***	-0.6129***
	(0.152)	*(0.150)*	*(0.150)*	*(0.150)*
Online (*Ref: No*)				
Yes	-0.2488	0.2356	-0.2585	0.4792**
	(0.207)	*(0.204)*	*(0.206)*	*(0.206)*
CT Band (*Ref: A / B / C*)				
D / E	0.1443	-0.0795	-0.0224	-0.2623
	(0.183)	*(0.179)*	*(0.180)*	*(0.184)*
F / G / H	-0.0960	-0.3414	0.0442	-0.0514
	(0.247)	*(0.245)*	*(0.243)*	*(0.251)*
Age (*Ref: 18 – 40*)				
41 - 60	0.1496	-0.3903**	0.5659***	0.4873**
	(0.191)	*(0.186)*	*(0.189)*	*(0.191)*
61 +	-0.4545	-0.5096*	0.6650**	0.7663**
	(0.311)	*(0.309)*	*(0.313)*	*(0.311)*
Gender (*Ref: Female*)				
Male	0.0514	0.1654	0.1944	-0.0742
	(0.152)	*(0.151)*	*(0.153)*	*(0.157)*
Employment (*Ref: Employed*)				
Unemployed	0.1341	0.0289	-0.1162	-0.3828
	(0.235)	*(0.227)*	*(0.246)*	*(0.241)*
Retired	0.8354***	-0.0616	-0.0897	-0.3510
	(0.289)	*(0.288)*	*(0.290)*	*(0.287)*
Degree (*Ref: No*)				
Yes	0.2775*	-0.1740	0.1496	-0.2497
	(0.163)	*(0.161)*	*(0.162)*	*(0.163)*
Location (*Ref: Urban*)				
Rural	-0.2488	-0.3381*	-0.1151	0.1709
	(0.181)	*(0.180)*	*(0.179)*	*(0.183)*
Bill Payer (*Ref: No*)				
Yes	-0.0434	-0.0307	0.4165*	-0.3162
	(0.240)	*(0.240)*	*(0.240)*	*(0.244)*
Bill Discount (*Ref: No*)				
Yes	0.0443	-0.2265	0.1645	0.1176
	(0.171)	*(0.168)*	*(0.174)*	*(0.173)*
Issue Experience (*Ref: Yes – Within 1 Year*)				
Yes - More than 1 Year ago	0.4325*	-0.1451	-------	-------
	(0.227)	*(0.317)*	-------	-------
No – Never	0.2874	0.0958	-------	-------
	(0.208)	*(0.244)*	-------	-------
Water Connection (*Ref: Public Supply*)				
Private Supply	-------	-------	-0.0543	-------
	-------	-------	*(0.411)*	-------
Fossil Fuel Concern (*Ref: Yes*)				
Neutral	-------	-------	-------	1.1316***
	-------	-------	-------	*(0.180)*
No	-------	-------	-------	1.7899***
	-------	-------	-------	*(0.216)*
Number of Participants	599	599	599	599

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.1. Results of ordered logistic models for the Ranking task with log odds reported (coefficients for “Prefer not to say” responses not reported)

Ranking Task	**Model 5** Supply Pipe Leakages	**Model 6** Drinking Water Quality	**Model 7** River/ Beach Cleanliness
Order – First Page (*Ref: No*)			
Yes	-0.6052***	-0.8422***	-0.7414***
	(0.149)	*(0.152)*	*(0.149)*
Online (*Ref: No*)			
Yes	-0.2468	-0.3046	0.1946
	(0.200)	*(0.208)*	*(0.211)*
CT Band (*Ref: A / B / C*)			
D / E	0.0723	0.0930	-0.0586
	(0.180)	*(0.176)*	*(0.179)*
F / G / H	0.0127	0.1815	-0.0535
	(0.247)	*(0.257)*	*(0.243)*
Age (*Ref: 18 – 40*)			
41 - 60	-0.5938***	-0.2979	0.0022
	(0.188)	*(0.191)*	*(0.190)*
61 +	-1.1120***	-0.3372	0.6720**
	(0.301)	*(0.300)*	*(0.311)*
Gender (*Ref: Female*)			
Male	-0.3342**	-0.2858*	0.1553
	(0.152)	*(0.151)*	*(0.153)*
Employment (*Ref: Employed*)			
Unemployed	-0.0086	0.1933	0.0886
	(0.230)	*(0.235)*	*(0.236)*
Retired	-0.1504	0.3166	-0.4921*
	(0.281)	*(0.281)*	*(0.290)*
Degree (*Ref: No*)			
Yes	-0.1100	0.3151*	-0.0352
	(0.164)	*(0.162)*	*(0.162)*
Location (*Ref: Urban*)			
Rural	-0.2619	0.2671	0.4526**
	(0.180)	*(0.177)*	*(0.179)*
Bill Payer (*Ref: No*)			
Yes	-0.4671**	0.4418*	0.1118
	(0.235)	*(0.242)*	*(0.240)*
Bill Discount (*Ref: No*)			
Yes	0.0391	0.0426	0.0056
	(0.167)	*(0.169)*	*(0.169)*
Water Site Visits (*Ref: More than Monthly*)			
Less than Monthly	-------	-------	0.1595
	-------	-------	*(0.153)*
Number of Participants	599	599	599

Table D.2. Results of ordered logistic models for the Ranking task with log odds reported (coefficients for "Prefer not to say" responses not reported)

<u>Appendix E</u> – Basic Slider Task vs. Cost Revelation Task Parametric Results

Here we report the results of statistical analyses which aim to isolate the individual effects of stage type, platform, experience and demographic effects on participants' allocations to each of the four *target improvements* in the Basic Slider and Cost Revelation tasks. Table E.1 reports the absolute differences in allocations across the Basic Slider and Cost Revelation tasks. Table E.2 reports the platform, experience and demographic effects on these differences between the two tasks. For ease of interpretation, Tables E.1 and E.2 do not report the coefficients for "I don't know"/ "Prefer not to say" responses to any experience or demographic questions, although these were controlled for in the models. Any meaningful effects of platform, demographic or experience effects in this stage are reported in Appendices H and J-L.

Shapiro-Wilk tests of normality for each *target improvement* suggest that the absolute allocations were not normally distributed in either the Basic Slider or Cost Revelation tasks for all *target improvement* (p< 0.001 for all tests) rendering any OLS regression inappropriate. As such the raw allocations were categorised into five categories: 0% - 5%; 10% - 15%; 20% - 25%; 30% - 35%; 40% +. Due to the non-independence of responses at the individual level a random-effects panel ordered logistic regression was used, clustered at the individual level, with allocation category as the dependent variable (although results did not substantively differ when a random-effects linear regression was run on the raw allocations).

In all models in Table E.1, the difference between the Basic Slider and Cost Revelation tasks mimics the results of the non-parametric statistics, once allowing for other demographic effects. Revealing cost information reduces likelihood to give a higher allocation for External Sewer Flooding and Rural Supplies, but results are not statistically significant (p= 0.295, p= 0.979, respectively), and significantly reduces likelihood to give a higher allocation for Renewable Energy (p= 0.058). Revealing cost information increases likelihood to give a higher allocation for Interruptions to Supply, but this is not statistically significant (p= 0.130).

When comparing the differences between the Basic Slider and Cost Revelation tasks, Shapiro-Wilk tests of normality for each *target improvement* suggest that the differences in allocations were not normally distributed (p< 0.001 for all tests) rendering any OLS regression inappropriate. As such the raw differences were categorised into five categories: Less than -10%; -5% - -10%; 0%; 5% - 10%; Greater than 10%. In all models in Table E.2 an ordered logistic regression was run, and these models satisfied the standard assumption checks for ordered logistic regression models[6].

[6] A Brant test for the proportional odds assumption could not be computed for all of the models as a result of the inclusion of "Prefer Not to Say" choices for demographic variables, where there was insufficient sample size to compute binary logits across all variables. Models 1 (p= 0.435), 2 (p= 0.335), 3 (p= 0.555) and 4 (p= 0.661) all passed the proportional odds assumption when these "Prefer Not to Say" choices were removed, without substantively changing results.

Basic Slider vs. Cost Revelation *Absolute Allocation*	**Model 1** Interruptions to Supply	**Model 2** External Sewer Flooding	**Model 3** Rural Supplies	**Model 4** Renewable Energy
Stage (*Ref: Basic Slider*)				
Cost Revelation	0.1825	-0.1220	-0.0031	-0.2330*
	(0.120)	*(0.117)*	*(0.119)*	*(0.123)*
Online (*Ref: No*)				
Yes	0.4761	0.2518	-0.2267	-0.6565
	(0.396)	*(0.304)*	*(0.389)*	*(0.422)*
CT Band (*Ref: A / B / C*)				
D / E	-0.1925	-0.2056	0.2942	0.0233
	(0.344)	*(0.264)*	*(0.337)*	*(0.369)*
F / G / H	-0.1260	0.4467	-0.1774	-0.4646
	(0.478)	*(0.373)*	*(0.468)*	*(0.506)*
Age (*Ref: 18 – 40*)				
41 - 60	1.2008***	0.9339***	-0.5210	-0.4435
	(0.367)	*(0.280)*	*(0.355)*	*(0.383)*
61 +	2.1671***	0.6721	-1.1633**	-0.8325
	(0.605)	*(0.459)*	*(0.582)*	*(0.621)*
Gender (*Ref: Female*)				
Male	0.3276	-0.1417	-0.5569*	-0.1308
	(0.293)	*(0.226)*	*(0.287)*	*(0.311)*
Employment (*Ref: Employed*)				
Unemployed	0.8734*	-0.3972	-0.6213	0.3015
	(0.455)	*(0.342)*	*(0.440)*	*(0.477)*
Retired	-1.2408**	0.7658*	0.4902	0.3152
	(0.558)	*(0.431)*	*(0.543)*	*(0.576)*
Degree (*Ref: No*)				
Yes	-0.5429*	0.1554	-0.6665**	0.4381
	(0.310)	*(0.240)*	*(0.305)*	*(0.331)*
Location (*Ref: Urban*)				
Rural	-0.6971**	0.1274	0.6297*	-1.2298***
	(0.349)	*(0.270)*	*(0.343)*	*(0.368)*
Bill Payer (*Ref: No*)				
Yes	0.0562	0.3215	-0.6957	-0.3247
	(0.456)	*(0.356)*	*(0.448)*	*(0.486)*
Bill Discount (*Ref: No*)				
Yes	0.1125	0.1401	-0.2127	-0.4264
	(0.326)	*(0.252)*	*(0.322)*	*(0.345)*
Issue Experience (*Ref: Yes – Within 1 Year*)				
Yes - More than 1 Year ago	-0.6986	0.0985	-------	-------
	(0.427)	*(0.499)*	-------	-------
No - Never	-0.3843	-0.0166	-------	-------
	(0.394)	*(0.384)*	-------	-------
Water Connection (*Ref: Public Supply*)				
Private Supply	-------	-------	-0.6745	-------
	-------	-------	*(0.733)*	-------
Fossil Fuel Concern (*Ref: Yes*)				
Neutral	-------	-------	-------	-2.1998***
	-------	-------	-------	*(0.370)*
No	-------	-------	-------	-3.9724***
	-------	-------	-------	*(0.447)*
Number of Observations	1,198	1,198	1,198	1,198
Number of Participants	599	599	599	599

Table E.1. Results of panel ordered logistic models for Basic Slider/ Cost Revelation task with log odds reported (coefficients for “Prefer not to say” responses not reported)

Basic Slider vs. Cost Revelation *Relative Difference*	**Model 1** Interruptions to Supply	**Model 2** External Sewer Flooding	**Model 3** Rural Supplies	**Model 4** Renewable Energy
Online (*Ref: No*)				
Yes	-0.0867	0.0041	0.2507	-0.0773
	(0.216)	*(0.218)*	*(0.221)*	*(0.223)*
CT Band (*Ref: A / B / C*)				
D / E	0.0468	-0.0462	0.1649	0.0106
	(0.186)	*(0.183)*	*(0.189)*	*(0.191)*
F / G / H	0.1048	-0.1226	-0.0523	0.2563
	(0.259)	*(0.262)*	*(0.264)*	*(0.262)*
Age (*Ref: 18 – 40*)				
41 - 60	0.1628	-0.2117	0.1897	0.2545
	(0.197)	*(0.195)*	*(0.196)*	*(0.197)*
61 +	0.6267*	-0.1201	-0.1505	-0.0291
	(0.323)	*(0.318)*	*(0.325)*	*(0.323)*
Gender (*Ref: Female*)				
Male	-0.3649**	0.2471	0.0427	0.2421
	(0.160)	*(0.157)*	*(0.160)*	*(0.159)*
Employment (*Ref: Employed*)				
Unemployed	0.3081	0.1485	-0.2672	-0.1401
	(0.245)	*(0.243)*	*(0.247)*	*(0.248)*
Retired	-0.3229	0.4608	0.0142	-0.2544
	(0.298)	*(0.295)*	*(0.303)*	*(0.299)*
Degree (*Ref: No*)				
Yes	0.1980	0.1458	-0.2134	0.0150
	(0.167)	*(0.167)*	*(0.170)*	*(0.170)*
Location (*Ref: Urban*)				
Rural	-0.0297	-0.0517	-0.5233***	0.3217*
	(0.187)	*(0.188)*	*(0.192)*	*(0.187)*
Bill Payer (*Ref: No*)				
Yes	0.5107**	-0.5648**	-0.0385	0.2021
	(0.248)	*(0.250)*	*(0.249)*	*(0.250)*
Bill Discount (*Ref: No*)				
Yes	-0.0041	0.2004	-0.3220*	-0.1413
	(0.176)	*(0.175)*	*(0.179)*	*(0.176)*
Issue Experience (*Ref: Yes – Within 1 Year*)				
Yes - More than 1 Year ago	-0.3020	0.0159	-------	-------
	(0.232)	*(0.354)*	-------	-------
No - Never	-0.1276	-0.2307	-------	-------
	(0.216)	*(0.279)*	-------	-------
Water Connection (*Ref: Public Supply*)				
Private Supply	-------	-------	0.5530	-------
	-------	-------	*(0.402)*	-------
Fossil Fuel Concern (*Ref: Yes*)				
Neutral	-------	-------	-------	-0.1877
	-------	-------	-------	*(0.183)*
No	-------	-------	-------	-0.2820
	-------	-------	-------	*(0.213)*
Number of Participants	599	599	599	599

Table E.2. Results of ordered logistic models for Basic Slider/ Cost Revelation task with log odds reported (coefficients for “Prefer not to say” responses not reported)

<u>Appendix F</u> - Cost Revelation vs. Time Horizon Task Parametric Results

Here we report the results of statistical analyses which aim to isolate the individual effects of stage type, platform and demographic effects on participants' allocations to each of the four time horizons in the Cost Revelation and Time Horizon tasks. Table F.1 reports the absolute differences in allocations across the Cost Revelation and Time Horizon tasks. Table F.2 reports the platform and demographic effects on these differences between the two tasks. For ease of interpretation, Tables F.1 and F.2 do not report the coefficients for "I don't know"/ "Prefer not to say" responses to any experience or demographic questions, although these were controlled for in the models. Any meaningful effects of platform or demographic in this stage are reported in Appendices H, J and K.

Shapiro-Wilk tests of normality for each time horizon suggest that the absolute allocations were not normally distributed in either the Cost Revelation or Time Horizon tasks for all time horizons ($p < 0.001$ for all tests) rendering any OLS regression inappropriate. As such the raw allocations were categorised into five categories: 0% - 5%; 10% - 15%; 20% - 25%; 30% - 35%; 40% +. Due to the non-independence of responses at the individual level a random-effects panel ordered logistic regression was used, clustered at the individual level, with allocation category as the dependent variable (although results did not substantively differ when a random-effects linear regression was run on the raw allocations).

In all models in Table F.1, the difference between the Cost Revelation and Time Horizon tasks mimics the results of the non-parametric statistics, once allowing for other demographic effects. Revealing time horizon information significantly increases likelihood to give a higher allocation for "Immediately" improvements ($p = 0.004$). It significantly reduces likelihood to give a higher allocation for "5 Years" improvements ($p = 0.024$), and produces no significant difference in likelihood to give a higher allocation for "10 Years" ($p = 0.648$) and "25 Years" ($p = 0.120$) improvements. For "Immediately" and "5 Years", higher allocations were more likely for External Sewer Flooding than Interruptions to Supply ($p < 0.001$ for both) and for "10 Years" and "25 Years", higher allocations were more likely for Renewable Energy than Rural Supplies ($p < 0.001$, $p = 0.009$, respectively).

When comparing the differences between the Cost Revelation and Time Horizon tasks, Shapiro-Wilk tests of normality for each time horizon suggest that the differences in allocations were not normally distributed ($p < 0.001$ for all tests) rendering any OLS regression inappropriate. As such the raw differences were categorised into five categories: Less than -10%; -5% - -10%; 0%; 5% - 10%; Greater than 10%. In all models in Table F.2 an ordered logistic regression was run, although not all models satisfied the standard assumption checks for ordered logistic regression models. However, the use of more appropriate models does not substantively change outcomes[7].

[7] A Brant test for the proportional odds assumption could not be computed for all of the models as a result of the inclusion of "Prefer Not to Say" choices for demographic variables, where there was insufficient sample size to compute binary logits across all variables. Models 1 ($p = 0.154$) and 3 ($p = 0.315$) passed the proportional odds assumption when these "Prefer Not to Say" choices were removed, without substantively changing results. Model 4 did not pass the proportional odds assumption when these "Prefer Not to Say" choices were removed ($p = 0.073$) but the sample size is sufficiently large and the difference in changes between categories does not influence interpretation of the key results, and so ordered logistic regression is a preferred model to a generalised ordered logistic regression. There were insufficient observations at the extreme categories for Model

Cost Revelation vs. Time Horizon *Absolute Allocation*	**Model 1** Immediately	**Model 2** 5 Years	**Model 3** 10 Years	**Model 4** 25 Years
Stage (*Ref: Cost Revelation*)				
Time Horizon	0.3579***	-0.2805**	0.0567	-0.1950
	(0.123)	*(0.124)*	*(0.124)*	*(0.125)*
Improvement (*Ref: Interruption to Supply*)				
External Sewer Flooding	1.4898***	1.3426***	-------	-------
	(0.306)	*(0.311)*	-------	-------
Improvement (*Ref: Rural Supplies*)				
Renewable Energy	-------	-------	1.4769***	0.9568***
	-------	-------	*(0.349)*	*(0.364)*
Online (*Ref: No*)				
Yes	0.0916	0.4237	-0.1263	-0.6670
	(0.417)	*(0.424)*	*(0.477)*	*(0.508)*
CT Band (*Ref: A / B / C*)				
D / E	0.1035	-0.4715	0.7995*	-1.0757**
	(0.365)	*(0.371)*	*(0.419)*	*(0.441)*
F / G / H	0.0124	0.1715	-0.0162	-1.0561*
	(0.509)	*(0.518)*	*(0.577)*	*(0.614)*
Age (*Ref: 18 – 40*)				
41 - 60	0.6091	1.5012***	-0.7613*	-0.1938
	(0.383)	*(0.394)*	*(0.439)*	*(0.460)*
61 +	0.8490	2.0234***	-1.2887*	-1.1762
	(0.634)	*(0.647)*	*(0.720)*	*(0.758)*
Gender (*Ref: Female*)				
Male	0.0002	0.1494	-0.4144	-0.4599
	(0.309)	*(0.316)*	*(0.354)*	*(0.373)*
Employment (*Ref: Employed*)				
Unemployed	-0.2965	0.5598	0.3273	-0.5692
	(0.473)	*(0.483)*	*(0.544)*	*(0.572)*
Retired	0.1345	-0.3487	0.6629	0.5047
	(0.590)	*(0.602)*	*(0.668)*	*(0.708)*
Degree (*Ref: No*)				
Yes	-0.2887	-0.0260	-0.1757	0.3869
	(0.330)	*(0.337)*	*(0.379)*	*(0.399)*
Location (*Ref: Urban*)				
Rural	-0.0168	-0.4024	-0.7513*	-0.2372
	(0.368)	*(0.373)*	*(0.421)*	*(0.441)*
Bill Payer (*Ref: No*)				
Yes	0.3523	-0.3483	-0.6048	-0.2631
	(0.484)	*(0.493)*	*(0.558)*	*(0.585)*
Bill Discount (*Ref: No*)				
Yes	-0.0208	0.2891	-0.1126	-0.6611
	(0.347)	*(0.352)*	*(0.395)*	*(0.418)*
Number of Observations	1,198	1,198	1,198	1,198
Number of Participants	599	599	599	599

Table F.1. Results of panel ordered logistic models for Cost Revelation/ Time Horizon task with log odds reported (coefficients for "Prefer not to say" responses not reported)

2. Pooling categories 1 and 2, and categories 4 and 5 in Model 2 overcame this issue without substantively changing results and the model passed the proportional odds assumption (*p*= 0.565).

Cost Revelation vs. Time Horizon *Relative Difference*	**Model 1** Immediately	**Model 2** 5 Years	**Model 3** 10 Years	**Model 4** 25 Years
Improvement (*Ref: Interruption to Supply*)				
External Sewer Flooding	-0.0125	0.3532**	-------	-------
	(0.156)	*(0.158)*	-------	-------
Improvement (*Ref: Rural Supplies*)				
Renewable Energy	-------	-------	-0.0328	-0.0369
	-------	-------	*(0.158)*	*(0.158)*
Online (*Ref: No*)				
Yes	-0.2632	0.1277	-0.0038	0.1590
	(0.212)	*(0.215)*	*(0.218)*	*(0.219)*
CT Band (*Ref: A / B / C*)				
D / E	0.2857	-0.1324	0.0828	-0.2681
	(0.190)	*(0.189)*	*(0.190)*	*(0.194)*
F / G / H	0.1303	0.4272	-0.1095	-0.3084
	(0.261)	*(0.264)*	*(0.266)*	*(0.265)*
Age (*Ref: 18 – 40*)				
41 - 60	-0.1094	-0.0537	0.0767	0.0785
	(0.199)	*(0.199)*	*(0.201)*	*(0.202)*
61 +	-0.6541*	-0.2104	0.3215	0.2087
	(0.335)	*(0.321)*	*(0.334)*	*(0.325)*
Gender (*Ref: Female*)				
Male	0.0900	-0.0416	0.0750	-0.2312
	(0.160)	*(0.161)*	*(0.163)*	*(0.162)*
Employment (*Ref: Employed*)				
Unemployed	-0.0402	-0.6923***	0.1067	0.7299***
	(0.249)	*(0.247)*	*(0.247)*	*(0.253)*
Retired	0.3140	-0.3193	-0.1022	0.1605
	(0.309)	*(0.296)*	*(0.311)*	*(0.302)*
Degree (*Ref: No*)				
Yes	-0.1746	-0.0034	0.0216	0.0906
	(0.171)	*(0.173)*	*(0.173)*	*(0.174)*
Location (*Ref: Urban*)				
Rural	-0.0820	0.0993	0.2571	-0.2035
	(0.189)	*(0.193)*	*(0.192)*	*(0.191)*
Bill Payer (*Ref: No*)				
Yes	-0.2949	0.0332	-0.2820	0.5788**
	(0.253)	*(0.256)*	*(0.252)*	*(0.254)*
Bill Discount (*Ref: No*)				
Yes	-0.2559	0.0245	0.1421	-0.0897
	(0.181)	*(0.180)*	*(0.183)*	*(0.180)*
Number of Participants	599	599	599	599

Table F.2. Results of ordered logistic models for Cost Revelation/ Time Horizon task with log odds reported (coefficients for “Prefer not to say” responses not reported)

<u>Appendix G</u> - Willingness-to-Pay Tasks Parametric Results

Here we report the results of statistical analyses which aim to isolate the individual effects of stage type, platform, experience and demographic effects on participants' willingness-to-pay. Table G.1 reports the results of four models – Models 1-3 report the results of logistic regressions, with the likelihood a participant would increase their willingness-to-pay by any amount (i.e. above £0) for each subsequent slider task as the dependent variable. Model 4 reports the change in willingness-to-pay by stage type, platform, experience and demographic effects, conditional on participants indicating willingness to pay some increase in each stage. Due to the extreme right-skew of responses to the willingness-to-pay stage, in Model 4 the dependent variable is the logarithmic transformation of willingness-to-pay, and a Tobit regression was deemed appropriate to account for the lower censoring of the data. In addition, to account for the non-independence of responses at the individual level in Model 4 a random-effects panel Tobit regression was used, clustered at the individual level. As a robustness check, Model 5 reports the results of an ordered logistic regression, when categorising willingness-to-pay values according to the categories found in Table 3 in the main text. To allow for sufficient variation across stages and categories of willingness-to-pay, Model 5 excludes participants who did not report any change in willingness-to-pay across the three stages. Table G.1 does not report the coefficients for "I don't know"/ "Prefer not to say" responses to any experience or demographic questions, although these were controlled for in the models. Any meaningful effects of platform or demographic effects in this stage are reported in Appendices H, J, and K.

Results from Model 1 infer that platform type (being a face-to-face participant) ($p < 0.001$) and having a degree ($p = 0.005$) substantially increase the likelihood a participant would increase their willingness-to-pay without any additional cost or time horizon information. Face-to-face ($p = 0.015$) and younger participants, compared to 41-60 ($p = 0.010$) and 61+ ($p = 0.026$) year olds, were significantly more likely to increase their willingness-to-pay further once provided cost information, as seen in Model 2. Effects are reduced in Model 3, suggesting demographic characteristics were less significant in influencing the likelihood that a participant would increase their willingness-to-pay further once provided time horizon information, although this may in part be a result of the relatively small number of participants overall who increased their willingness-to-pay further in this task.

The results from Model 4 support the findings of the non-parametric statistics. Compared to the Cost Revelation task, participants had a significantly lower willingness-to-pay in the first Willingness-to-Pay task ($p < 0.001$) and significantly higher willingness-to-pay in the Time Horizon task ($p = 0.025$). In addition, and in support of the non-parametric statistics, there was a positive and increasing effect of "Top Four Ranking Sum"; the number of *target improvements* ranked 1st – 4th in the Ranking task. Compared to those with only 1 *target improvements* ranked 1st – 4th in the Ranking task, those with 2 ($p = 0.007$) and 3 ($p = 0.022$) had a significantly higher willingness-to-pay overall. The coefficients and standard errors imply that there would not be a statistically significant difference between 2 and 3. The willingness-to-pay of online participants was significantly lower than those face-to-face ($p = 0.002$).

The results from Model 5 support the main findings of Model 4. Compared to the Cost Revelation task, participants were significantly less likely to report a higher category of willingness-to-pay in the first Willingness-to-Pay task ($p < 0.001$) and significantly more likely to report a higher category of willingness-to-pay in the Time Horizon task (p= 0.009). The effects of platform type (p= 0.203) and “Top Four Ranking Sum” (p= 0.527, p= 0.134) were no longer significant. One plausible reason for this is that much of the overall effects of platform and “Top Four Ranking Sum” are being driven by participants who do not report differences in willingness-to-pay across tasks.

Willingness-to-Pay	**Model 1** Logit **WTP**	**Model 2** Logit **Cost Revelation**	**Model 3** Logit **Time Horizon**	**Model 4** Tobit (Log WTP) **All**	**Model 5** Ordered Logit **All**
Stage (*Ref: Cost Revelation*)					
Willingness-to-Pay	-------	-------	-------	-0.0912***	-2.1019***
	-------	-------	-------	*(0.014)*	*(0.291)*
Time Horizon	-------	-------	-------	0.0314**	0.6681***
	-------	-------	-------	*(0.014)*	*(0.254)*
Online (*Ref: No*)					
Yes	-1.7319***	-0.7346**	-0.1039	-0.4441***	-1.4315
	(0.325)	*(0.302)*	*(0.440)*	*(0.145)*	*(1.125)*
Top Four Rating Sum (*Ref: One*)					
Two	0.3380	-0.4808*	-0.5303	0.4176***	0.6976
	(0.223)	*(0.285)*	*(0.352)*	*(0.156)*	*(1.103)*
Three	0.0635	-0.4397	-2.1298***	0.4419**	2.2653
	(0.284)	*(0.370)*	*(0.774)*	*(0.193)*	*(1.511)*
Four	-------	-------	-------	-------	-------
	-------	-------	-------	-------	-------
CT (*Ref: A/B/C*)					
D / E	-0.2344	0.0404	-0.1405	0.2893**	0.5247
	(0.220)	*(0.295)*	*(0.406)*	*(0.140)*	*(1.182)*
F / G / H	-0.3206	-0.1906	-0.5641	0.2202	1.1320
	(0.310)	*(0.433)*	*(0.632)*	*(0.190)*	*(1.686)*
Age (*Ref: 18 – 40*)					
41 - 60	-0.0930	-0.8072***	-0.9481**	-0.0159	-0.6330
	(0.233)	*(0.312)*	*(0.445)*	*(0.151)*	*(1.252)*
61 +	0.3911	-1.1790**	-0.0084	-0.0142	1.0566
	(0.385)	*(0.528)*	*(0.630)*	*(0.216)*	*(1.887)*
Gender (*Ref: Female*)					
Male	-0.3340*	-0.0235	0.3140	-0.0940	-1.0516
	(0.187)	*(0.248)*	*(0.340)*	*(0.124)*	*(0.971)*
Employment (*Ref: Employed*)					
Unemployed	0.5085*	-0.1846	0.5369	-0.1626	-0.4618
	(0.297)	*(0.401)*	*(0.470)*	*(0.177)*	*(1.505)*
Retired	-0.0608	0.6848	-0.1797	-0.2186	-2.7814
	(0.353)	*(0.493)*	*(0.629)*	*(0.206)*	*(1.750)*
Degree (*Ref: No*)					
Yes	0.5710***	-0.1170	0.2754	-0.0995	-1.0542
	(0.202)	*(0.269)*	*(0.363)*	*(0.128)*	*(1.090)*
Location (*Ref: Urban*)					
Rural	-0.2589	0.2167	-0.9309*	-0.0680	0.3745
	(0.215)	*(0.293)*	*(0.506)*	*(0.150)*	*(1.218)*
Bill Payer (*Ref: No*)					
Yes	0.0837	0.6908*	0.2805	0.0276	-0.4759
	(0.300)	*(0.419)*	*(0.567)*	*(0.190)*	*(1.531)*
Bill Discount (*Ref: No*)					
Yes	-0.2261	-0.3406	0.0278	-0.0208	-1.4114
	(0.208)	*(0.280)*	*(0.369)*	*(0.139)*	*(1.103)*
Constant	1.8534***	-0.9034*	-1.9854***	2.5993***	-------
	(0.464)	*(0.539)*	*(0.741)*	*(0.259)*	-------
Number of Observations	596	590	590	1,062	375
Number of Participants	-------	-------	-------	354	125

Table G.1. Results of logistic regression models (Models 1-3), random-effects panel Tobit regression model (Model 4) and ordered logistic model for willingness-to-pay tasks with log odds reported in Models 1-3 and 5 (coefficients for "Prefer not to say" responses not reported)

<u>Appendix H</u> - Online vs. Face-to-Face

This study was conducted both online and face-to-face. The degree to which responses differ across these platforms may raise methodological questions as to the role of online or laboratory studies in eliciting the attitudes of citizens. In general, the demographic characteristics between the two platform types were similar. Given our categorisation of demographic information, online and face-to-face participants did not differ by council tax group (χ^2= 4.523, *p*= 0.104) age (χ^2= 4.372, *p*= 0.112), gender (χ^2= 0.042, *p*= 0.839), employment type (χ^2= 0.752, *p*= 0.687), educational attainment (χ^2= 1.525, *p*= 0.217) or council tax discount receipt (χ^2= 0.401, *p*= 0.526). A greater proportion of online participants lived in rural (rather than urban) locations (χ^2= 5.289, *p*= 0.021) and were the main water bill payers (χ^2= 3.342, *p*= 0.068), relative to face-to-face participants.

<u>Ranking Task</u>

Table H.1 below outlines the average rankings for each improvement when separated by study type.

	Average Ranking (/ 7)			
	Online	**Face-to-Face**	***t*-stat**	***p*-value**
Interruptions to Supply	4.73	5.10	1.781	0.075
External Sewer Flooding	3.78	3.54	-1.234	0.218
Rural Supplies	4.60	4.87	1.288	0.198
Renewable Energy	4.52	3.84	-2.981	0.003
Target Improvements	**4.41**	**4.34**	**-1.001**	**0.318**
Supply Pipe Leakages	3.34	3.57	1.139	0.255
Drinking Water Quality	3.75	4.04	1.224	0.222
Beach/ River Cleanliness	3.28	3.05	-1.169	0.243
Other Improvements	**3.46**	**3.55**	**1.001**	**0.318**

Table H.1. Average ranking of improvements by platform type

For the majority of improvements there are few substantial differences between platform types. However, there is a large and statistically significant difference in the preferences for Renewable Energy, where face-to-face respondents ranked this improvement as significantly more important (*t*= -2.981, *p*= 0.003). The reverse was true for Interruptions to Supply, where online respondents ranked this improvement as significantly more important, although not to the same magnitude (*t*= 1.781, *p*= 0.075). Including platform type in the parametric models of each improvement in the Ranking task, platform type only significantly affected Renewable Energy, where online participants were significantly less likely to give this improvement a lower ranking, (*p*= 0.020), as shown in Tables D.1 and D.2.

Slider Tasks

Table H.2 below reports the average allocations in the Basic Slider task between online and face-to-face participants.

	Average Allocation			
	Online	**Face-to-Face**	***t*-stat**	***p*-value**
Interruptions to Supply	22.8%	20.6%	-1.735	0.083
External Sewer Flooding	29.0%	26.9%	-1.517	0.130
Personal Improvements	**51.9%**	**47.5%**	**-2.553**	**0.011**
Rural Supplies	21.6%	23.0%	1.025	0.306
Renewable Energy	26.6%	29.5%	1.760	0.079
Societal Improvements	**48.2%**	**52.5%**	**2.553**	**0.011**

Table H.2. Average allocations in the Basic Slider task by platform type

In keeping with the overall consistency between Ranking task and Basic Slider task as found in the main text, online participants allocated significantly more than face-to-face participants to Interruptions to Supply (t= -1.735, p= 0.083), and significantly less to Renewable Energy (t= 1.760, p= 0.079). Combining improvement types, online participants gave significantly more to *personal* improvements (and necessarily equivalently, significantly less to *societal* improvements) than face-to-face participants (t= -2.553, p= 0.011).

Table H.3 below reports the average allocation differences between Basic Slider and Cost Revelation tasks, and between Cost Revelation and Time Horizon tasks. There were no significant differences in average difference as a result of additional information revelation. This suggests that, whilst participants on different platforms revealed systematic differences in their absolute preference for different types of improvements, there was no evidence of differences in the relative effects of information provision about these specific improvements. A lack of statistically significant differences between allocations is reflected in the parametric statistics in Tables E.1 and E.2, and Tables F.1 and F.2.

	Average Allocation – Basic Slider and Cost Revelation			
	Online	**Face-to-Face**	***t*-stat**	***p*-value**
Interruptions to Supply	0.7%	0.6%	-0.049	0.961
External Sewer Flooding	-0.5%	0.9%	1.088	0.277
Rural Supplies	0.4%	-1.4%	-1.496	0.135
Renewable Energy	-0.6%	-0.2%	0.347	0.729
	Average Allocation – Cost Revelation and Time Horizon			
	Online	**Face-to-Face**	***t*-stat**	***p*-value**
Immediately	1.3%	1.6%	0.245	0.807
5 Years	-0.9%	-0.7%	0.207	0.836
10 Years	-0.2%	0.2%	0.335	0.738
25 Years	-0.3%	-1.1%	-0.897	0.370

Table H.3. Average differences in allocation across slider tasks by platform type

Willingness-to-Pay Tasks

Table H.4 below reports the average willingness-to-pay and changes across slider tasks by platform type, and Table H.5 below reports the distribution of responses in the Willingness-to-Pay task, separated by platform type. There is a large and statistically significant difference between the average willingness-to-pay by platform type, with face-to-face participants overall willing to pay 112.5% more than online participants (t= 5.423, p< 0.001). There is a substantial difference in the number of participants who would not increase their current annual water charge. Online participants were more than three times more likely to indicate a willingness-to-pay additional charges of £0.00 than face-to-face participants (χ^2= 33.20, p< 0.001). However, even amongst those willing to add something to their water charges, face-to-face participants were willing to pay 35.5% more than online participants (t= 2.443, p= 0.015). There was no statistically significant difference between platform type in the extent to which overall willingness-to-pay was influenced by either the Cost Revelation task (t= -0.023, p= 0.982) or the Time Horizon task (t= 0.217, p= 0.829). A smaller proportion of online participants increased their willingness-to-pay when costs were revealed compared to face-to-face (t= 5.367, p= 0.021). There was no difference in the proportion of respondents who increased their willingness-to-pay when time horizons were revealed (t= 0.425, p= 0.514). These differences are reflected in the parametric analyses of Table G.1.

	Willingness-to-Pay			
	Online	**Face-to-Face**	***t*-stat**	***p*-value**
All	£9.30	£19.76	5.423	< 0.001
WTP > £0.00	£16.98	£23.01	2.443	0.015
	Online	**Face-to-Face**	**χ^2-stat**	***p*-value**
WTP = £0.00	45.2%	14.1%	33.20	< 0.001
	Average Difference – Cost Revelation and Willingness-to-Pay			
	Online	**Face-to-Face**	***t*-stat**	***p*-value**
All	+ £1.06	+ £1.05	-0.023	0.982
No WTP Increase	86.8%	77.8%	5.367	0.021
	Average Difference – Time Horizon and Cost Revelation			
	Online	**Face-to-Face**	***t*-stat**	***p*-value**
All	+ £0.46	+ £0.54	0.217	0.829
No WTP Increase	92.8%	90.1%	0.425	0.514

Table H.4. Difference in willingness-to-pay responses by platform type

	Online		Face-to-Face	
	(n)	(%)	(n)	(%)
£0	226	45.2%	14	14.1%
£1 - £5	83	16.6%	14	14.1%
£6 - £10	63	12.6%	11	11.1%
£11 - £20	70	14.0%	25	25.3%
£21 - £30	23	4.6%	15	15.2%
£31 - £40	8	1.6%	6	6.1%
£41 - £50	10	2.0%	7	7.1%
£51 - £75	13	2.6%	7	7.1%
£76 +	4	0.8%	0	0.0%

Table H.5. Distribution of willingness-to-pay responses in Willingness-to-Pay task, by platform type

Together, the results separated by platform type imply a consistent theme. Although there are systematic differences between responses by platform, these are typically across absolute decisions (e.g. initial rankings, allocations or willingness-to-pays). There is little overall evidence of systematic differences in the relative decisions as a result of experimental changes. This suggests that participants are not responding differently to the changing stimuli within the experiment but rather have different a-priori preferences that differ according to platform type.

It is not straightforward to identify the cause of these systematic differences. Face-to-face participants were paid substantially more for taking part than online participants, although there are potentially greater costs of attending an experimental session face-to-face. It could be plausible that a greater initial willingness-to-pay for face-to-face respondents is a reciprocal response to a higher participation payment. However, the finding of greater overall allocation to *societal* improvements over *personal* improvements is independent of any payment decisions. This suggests that the type of individual who engages with face-to-face market research studies is of a more altruistic predisposition than those who participate in online market research studies.

Efforts have been made to minimise the potential for *experimenter demand effects* (e.g. Zizzo, 2010) in this study. For most plausible issues, there would be little reason to distinguish these across platform type. However, face-to-face participants conducted the experiment in groups of approximately ten per session. The physical clicking of the mouse required to indicate an increased willingness-to-pay might have increased a social pressure to increase one's own willingness-to-pay in the face-to-face studies that would not have been present in the online study. Again, this would not explain a propensity to allocate more to *societal* improvements. An alternative argument could be made that the mere presence of experimenters in the room whilst face-to-face participants were completing the study would encourage participants to indicate a greater willingness-to-pay. However, given that *societal* improvements were not marked as such, and were randomly ordered, there would have been no explicit indication that allocating more to *societal* improvements would in any way be beneficial to the experimenters (even if the experimenters were perceived to be representatives of the Scottish water industry). Altogether, the differences between platform types suggest a genuine difference in the a-priori attitudes of online and face-to-face respondents.

<u>Appendix I</u> - Order Effects

In the Ranking task the order in which participants saw the descriptions of improvements was randomised across participants. In the Basic Slider task (and all subsequent slider tasks) the order in which participants saw the sliders on their screens was randomised across participants. From a methodological perspective it is important to consider whether the order in which participants saw improvements and sliders influenced their decisions.

<u>Ranking Task</u>

In the Ranking task, participants first saw the seven descriptions of possible water service issues and a potential improvement for each. These were separated over two pages, with four shown on the first screen and three on the second. The order of the specific descriptions was randomised across participants. The order of the descriptions was the same as the order of the ranking options in the ranking task. Therefore any participants who were not fully engaging with the ranking task (and instead simply ranking the improvements in the order in which they appeared in a list) would appear to be exhibiting order effects. There were three such participants, and these three are removed from subsequent analyses.

Table I.1 below outlines the average ranking and likelihood to be ranked most important, by order of presentation.

	Average Ranking (/ 7)	**Proportion of Ranking**	
		1st	**1st - 4th**
1st	3.24	29.5%	70.3%
2nd	3.59	20.8%	65.8%
3rd	3.93	13.4%	59.7%
4th	4.03	11.4%	55.5%
First Page	**3.69**	**18.8%**	**62.8%**
5th	4.28	7.9%	52.7%
6th	4.48	8.7%	48.2%
7th	4.46	8.2%	47.8%
Second Page	**4.41**	**8.3%**	**49.6%**

Table I.1. Average ranking and likelihood to be ranked as 1st or 1st - 4th most important by presentation order

There is strong evidence of an order effect on improvement ranking. Those that were shown on the first page were ranked lower (i.e. more important) on average than those shown on the second page. In addition, when accounting for order in the parametric models of each improvement in the Ranking task, being shown on the first page significantly increased the likelihood of a lower ranking for all seven improvements ($p< 0.001$ for all), as shown in Tables D.1 and D.2.

It is possible that the order effects are simply a result of participants not engaging with the ranking task and having a tendency to first select the options at the top of the list of options.

However, we argue that this is implausible. First, the level of consistency between Ranking tasks and slider tasks suggests that participants were acting on some consistent decision making process between each task, but one that was unlikely to be a result of simply memorising the order of preference in the Ranking task and replicating this in the slider tasks. It thus seems paradoxical that participants would not engage with the Ranking task but then successfully replicate their preferences from this task in the later slider tasks. Second, response time data in the Ranking task can be used as a proxy for effort level in this task. One would anticipate a participant who did not engage fully (i.e. ranking the options in the order in which they were presented) would take less time than those who deliberated their options. Indeed, the three online participants who selected the improvements in the exact order in which they were displayed completed the task in approximately one third of the time of the average response time. The strongest evidence of order effects in this task is the occurrence of the first presented improvement to be ranked as most important. There was no statistically significant difference in the average response times of those who ranked the first presented improvement as most important (65.7 secs, *SD*: 53.35) and those who did not rank the first presented improvement as most important (59.1 secs, *SD*: 68.35) (t= -1.131, p= 0.259) with those who ranked the first presented improvement as most important actually taking longer in this task.

This finding is not problematic for the core aims of the study but does raise an interesting issue. It is probable that citizens do not frequently think about the range of potential water service improvements that the water industry may make. These results suggest that participants give disproportionate weight to the improvements they read about first. Given the level of consistency between the Ranking task and later slider tasks, this suggests that this weighting influenced not only their ranking of the improvements but also their consequent slider allocations and willingness-to-pay towards improvements. The findings here suggest that the order in which information is provided, and when, can have significant influence on how important a citizen judges that information to be.

Slider Task

In the Basic Slider task (and all consequent slider tasks), the order in which the four improvements were presented as sliders was randomised across participants. An order effect in these tasks might manifest itself as a greater increase in allocation to sliders presented at the top of each task. Figure I.1 below reports the average allocation by slider order in the Basic Slider task.

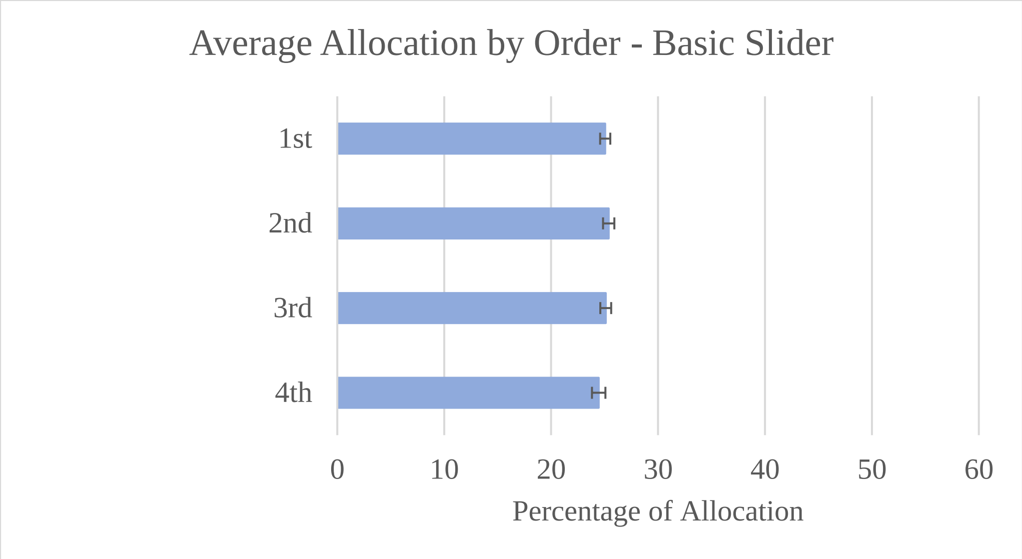


Figure I.1. Average allocation by slider order in Basic Slider task

As can be seen, there were no substantial differences in allocation by order, with allocations close to equal across orders, ranging from 24.4% (4th) to 25.4% (2nd). This suggests further still that the order effects found in the Ranking task were a result of the reading of descriptions rather than through disengagement with the task itself, since if order effects were a more general occurrence then they could be expected to be present in this task also.

Appendix J - Council Tax Band

As water charges in Scotland are fixed according to a property's council tax band, participants were asked to input their council tax band in order to provide them with an accurate current annual water charge from which price changes could be made.

As there were concerns that participants might not know their council tax band, a link to the Scottish Assessors Association website (www.saa.gov.uk) was provided, where participants could enter their address to find their own council tax band. Over one-fifth of participants (21.9%) availed of this option, suggesting that a substantial number of citizens are unaware (or required confirmation) of their council tax bands. An "I don't know" option was provided for participants who were still unsure of their council tax band. 9.5% of participants selected this option (it is of interest to note that none of the 57 participants who selected "I don't know" used the Scottish Assessors Association website link, suggesting that this was a useful tool for participants to find their council tax band).

Table J.1 below outlines the 2018/19 annual water charges in Scotland for each council tax band, as well as the proportion of the study participants who selected each band and the national distribution of council tax band across all registered properties in Scotland[8].

	Council Tax Band								
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**"I don't know"**
Annual Cost	£291.60	£340.20	£388.80	£437.40	£534.60	£631.80	£729.00	£874.80	N/A
Study (n)	59	91	89	121	97	46	36	3	57
Study (%)	10.9	16.8	16.4	22.3	17.9	8.5	6.6	0.6	N/A
National (%)	20.8	22.9	16.1	13.5	13.4	7.8	5.0	0.5	N/A

Table J.1. Distribution of council tax bands by study and national level

Relative to the national distribution, citizens with the lowest council tax bands were slightly underrepresented. This may suggest that price rise acceptance in this study could be an overestimation of the true population average. However, it is possible to measure whether differences exist between responses based on council tax band within this sample population.

Regarding the different stages of this study, there is some evidence that council tax band influenced responses in certain aspects of the study. Council tax bands were categorised into four groups (A/B/C; D/E; F/G/H; "I don't know"). There is some evidence that participants in higher council tax bands were less likely to allocate more to the "25 Years" improvement across both Cost Revelation and Time Horizon tasks, as shown in Table F.1 (D/E, p= 0.015, F/G/H, p= 0.086). However, the actual difference between allocations across the two tasks did not differ by council tax band, as shown in Table F.2. There is some evidence that participants in higher council tax bands had a higher willingness-to-pay across all tasks, although this was only statistically significant for participants in bands D/E (p= 0.038) and not F/G/H (p= 0.246),

[8] Taken from: Scottish Assessors Association, 2019, *Report 3 – Council Tax by Assessor/ Local Authority/ Council Tax Band*, [online] Scottish Assessors Association, Available at: <https:// https://www.saa.gov.uk/general-statistics/?REPORT_NAME=ct_band#report_list > [Accessed 24/04/2019]

as shown in Table G.1. Overall however, there were no large or systematic differences in responses across the whole study by council tax band.

Appendix K - Demographic Information

It is possible that belonging to certain demographic groups systematically influenced decisions throughout this study. Below we report any consistent and statistically significant differences in responses throughout this study.

Age: In the Ranking task, comparisons of each improvement in Tables D.1 and D.2 suggest there is a general tendency for older participants (i.e. those aged 41-60 and 61+, relative to participants aged 18-40) to be more likely to give lower rankings for *personal* improvements. We identify *personal* improvements to be Interruptions to Supply, External Sewer Flooding and Drinking Water Quality. The *societal* improvements for which participants may not themselves directly benefit were Rural Supplies, Renewable Energy, Supply Pipe Leakages and River/ Beach Cleanliness. The exception to this is Supply Pipe Leakages, where participants aged both 41-60 (p= 0.002) and 61+ (p< 0.001) give this improvement a lower ranking.

When pooling these improvements into *personal* and *societal*, older participants ranked *personal* improvements as significantly more important (41-60, p= 0.033; 61+, p= 0.097) when compared to 18-40 year olds. There was no significant difference between the two older groups (p= 0.664).The same was also true in the Basic Slider task, with older participants allocating significantly more to *societal* improvements than 18-40 year olds (41-60, p< 0.001; 61+, p< 0.001). There was no significant difference between the two older groups (p= 0.723). This is illustrated in Table K.1 below. A similar effect is found in comparisons of both Basic Slider and Cost Revelation in Table E.1, as well as Cost Revelation and Time Horizon in Table F.1 (where "Immediately" and "5 Years" feature only *personal* improvements).

	Average Ranking (/ 7)		**Basic Slider Allocation**	
	Personal	**Societal**	**Personal**	**Societal**
18 – 40	4.22	3.83	46.4%	53.6%
41 – 60	4.02	3.99	53.8%	46.2%
61 +	4.06	3.96	54.3%	45.7%
	***t*-stat**	***p*-value**	***t*-stat**	***p*-value**
18 – 40 vs. 41 – 60	2.146	0.033	4.916	< 0.001
18 – 40 vs. 61 +	1.662	0.097	5.168	< 0.001
41 – 60 vs. 61 +	-0.434	0.664	0.355	0.723

Table K.1. Comparison of improvement type by average ranking in the Ranking task and average allocation in the Basic Slider task

Gender: Overall there was little evidence of systematic differences in responses across the study by gender.

Employment: Compared to employed participants, retired participants in general indicated a reduced preference for Interruptions to Supply in both the Ranking task in Table D.1 (p= 0.004) and across the Basic Slider and Cost Revelation tasks in Table E.1 (p= 0.026). There were no other systematic differences in responses across the study by employment type.

Educational Attainment: Compared to participants without a degree, those with a degree in general indicated a reduced preference for Interruptions to Supply in both the Ranking task in

Table D.1 (p= 0.088) and across the Basic Slider and Cost Revelation tasks in Table E.1 (p= 0.080). In addition participants with a degree were significantly more willing to allow some price increase in their water charge in the initial willingness-to-pay task as shown in Table G.1 (p= 0.005). There was no significant difference by education attainment amongst those who were willing to increase their water charge, however.

Location: Participants living in rural locations allocated significantly more towards Rural Supplies across the Basic Slider and Cost Revelation tasks (p= 0.066) and significantly less towards Interruptions to Supply (p= 0.080) and Renewable Energy (p= 0.001), compared to participants living in urban locations, as shown in Table E.1. However, rural participants were less likely to give more to Rural Supplies once cost information was revealed than urban participants (p= 0.006) as shown in Table E.2.

Bill Payer/ Bill Discount: Overall there was little evidence of systematic differences in responses across the study by whether participants were the main bill payer of if they were in receipt of a council tax bill discount.

<u>Appendix L</u> - Experience Questionnaire

It was reasoned that the attitudes of participants throughout this study may in part be determined by their experience with the issues presented to them. Through a series of multiple choice questions, this stage aimed to elicit participants' experiences with the water industry in Scotland, and in particular their experiences with some of the specific issues presented throughout the study. The underlying premise was that participants who had experienced negative effects of service issues, or would be more directly affected by the negative consequences, would perceive these to be relatively more important than those who had never experienced them, or would not be so directly affected by their consequences.

The first question in this stage related to participants' experience of short term interruptions to supply. The response options asked participants if they had ever experienced short term interruptions to supply (lasting at least 6 hours) and if so to indicate how recently (via multiple choice response options ranging from "Yes, within the last month", to "Yes, longer than three years ago"). Overall, 51.8% of all participants indicated that they had experienced a short term interruption to supply in their homes at some point. Data from the Scottish water industry estimates approximately 5,250 households are affected by interruptions to supply of 6-12 hours each year (approximately 0.2% of all households in Scotland). 19.5% of participants in this study claimed to have experienced an interruption to supply of 6-12 hours within the last year. It is possible that participants did not correctly discern the length of time of the interruption to supply that they had experienced, and reported having experienced a short-term interruption to supply even if it was not as long as the 6 hour criteria. Alternatively, it is also possible that participants were reporting incidences of supply interruption outside of their own home (such as a neighbour, family member or in a work place).

Table L.1 below breaks down both the average ranking and basic slider allocation towards improving interruptions to supply by experience with this issue. 6.7% of participants selected the "I don't know" option. When comparing a straightforward binary split, there was no significant difference in average ranking (t= 0.679, p= 0.497) or basic slider allocation (t= 0.345, p= 0.730). There does appear to be a recency effect on relative importance, however, with those who experienced interruptions to supply within the last year ranking this improvement as significantly more important (t= 3.257, p= 0.001). This significant effect is observed in the parametric analysis of the Ranking task in Table D.1 also (p= 0.057). There was no significant difference by allocation amount (t= 0.766, p= 0.445).

Interruptions to Supply	**Ranking Task (/ 7)**	**Basic Slider Allocation**
Yes	4.80	22.5%
No	4.90	22.2%
Yes – Less than one year ago	4.34	23.2%
Yes – More than one year ago	5.07	22.1%
No	4.90	22.2%

Table L.1. Average ranking and basic slider allocation for Interruptions to Supply by experience of issue

The second question in this stage concerned experience with external sewer flooding, and was framed in the same way as the above question. Some 20.2% of participants claimed to have experienced external sewer flooding at some point. Table L.2 below breaks down both the average ranking and basic slider allocation towards improving external sewer flooding. 3.3% of participants selected the "I don't know" option. There were no significant differences in average ranking (t= 0.844, p= 0.399) or basic slider allocation (t= 0.688, p= 0.492) between those who had or had not experienced external sewer flooding, nor any evidence of a recency effect for average ranking (t=-1.066, p= 0.289) or basic slider allocation (t= -0.763, p= 0.447) for those who had.

External Sewer Flooding	Ranking Task (/ 7)	Basic Slider Allocation
Yes	3.60	28.0%
No	3.75	28.9%
Yes – Less than one year ago	3.75	27.0%
Yes – More than one year ago	3.45	28.8%
No	3.75	28.9%

Table L.2. Average ranking and basic slider allocation for External Sewer Flooding by experience of issue

The third question considered the type of water supply participants had in their homes. Those who are connected to a private water supply would directly benefit from increased investment in connecting to the public water system. Data from the Scottish water industry suggests approximately 3.6% of homes in Scotland have private water connections. 4.0% of participants indicated that they were connected to a private supply. 5.0% said they were connected to something other than the public/ private supply, or did not know.

Table L.3 below breaks down both the average ranking and basic slider allocation towards increasing rural connections by water supply type. There were no significant differences in average ranking (t= 0.661, p= 0.509) or basic slider allocation (t= 0.209, p= 0.834) by water supply type.

Rural Supplies	Ranking Task (/ 7)	Basic Slider Allocation
Public	4.67	21.8%
Private	4.42	21.3%

Table L.3. Average ranking and basic slider allocation for rural connections by water supply type

The fourth question aimed to elicit participants' general attitudes towards the reliance on fossil fuels and their impact on the environment. This was intended to be a proxy for general environmental awareness. How much an individual cares about the environment could impact the importance they give to increasing renewable energy reliance. The question participants were asked was "*How concerned are you of the effect of the use of fossil fuels on the environment in general (i.e. not just on the water system)?*" and this was adapted from Kelly *et al*. (2004). Responses ranged from "*extremely concerned*" to "*not at all concerned*", with "*somewhat concerned*" as a mid-point option. Table L.4 below breaks down both the average ranking and basic slider allocation towards increasing renewable energy production, by attitudes towards fossil fuel usage on the environment, categorising responses as positive "*extremely/ very concerned*", neutral "*somewhat concerned*" and negative "*not very/ not at all concerned*". 31.7% responded positively, 41.7% responded neutrally and 22.3% responded negatively. 4.2% of participants indicated that they were not sure or preferred not to say.

Those who responded positively ranked increased renewable energy production as significantly more important (t= 6.702, p< 0.001) and allocated significantly more investment in the basic slider task than those who responded neutrally (t= 6.537, p< 0.001). In turn, those who responded neutrally ranked increased renewable energy production as significantly more important (t= 3.667, p< 0.001) and allocated significantly more investment in the basic slider task than those who responded negatively (t= 4.351, p< 0.001). These effects are observed in the parametric analyses of the Ranking task in Table D.1 (p< 0.001 for both) and the Basic Slider and Cost Revelation tasks in Table E.1 (p< 0.001 for both).

Renewable Energy	**Ranking Task (/ 7)**	**Basic Slider Allocation**
Positive	3.36	34.8%
Neutral	4.64	25.6%
Negative	5.40	19.3%

Table L.4. Average ranking and basic slider allocation for increasing share of renewable energy by attitudes towards fossil fuel usage on the environment

Lastly, participants were also asked about the frequency of their visitations to water leisure sites (such as rivers, beaches and lochs). How important the cleanliness of beaches/ rivers would be may be dependent on how frequently a citizen visits them. Responses ranged from "*At least once a day*" to "*less than once a year*". 45.7% of participants reported visiting these sites at least once a month, and 50.4% reported visiting less than once a month or less. 3.8% of participants were not sure.

Table L.5 below breaks down average ranking of improving river/ beach cleanliness by water site visitation frequency. There was no significant difference in average ranking by visitation frequency (t= 0.826, p= 0.409).

Improving River/ Beach Cleanliness	**Ranking Task (/ 7)**
At least once a month	3.16
Less than once a month	3.28

Table L.5. Average ranking for improving river/ beach cleanliness by water site visitation frequency

References

Kelly, M., Kennedy, F., Faughnan, P., and Tovey, H. (2004), *Environmental attitudes and behaviours: Ireland in comparative European perspective*, Environmental Protection Agency

Williams, R. (2016), Understanding and interpreting generalized ordered logit models, *Journal of Mathematical Sociology*, *40*, 7-20

Zizzo, D. J. (2010), Experimenter demand effects in economic experiments, *Experimental Economics*, *13*, 75-98