

Newcastle, 15 February 2023

Estimating Customers' Willingness to Pay for Changes in Service at PR24

This report provides a peer review of the document "Estimating Customers' Willingness to Pay for Changes in Service at PR24" prepared by NERA for Wessex Water, in particular with reference to Ofwat's standards of high-quality research and customer challenge and engagement.

The document consists of 93 pages. Of these, nine pages constitute the Executive Summary and eight pages constitute the Appendices. The content of the Appendices C and D is listed but included in separate documents. The remaining 74 pages include the description of the entire work conducted.

Overall, this is a high-quality research, which fulfils the Ofwat's standards in terms of customer challenge and engagement. The document includes a section 2.8.1 specifically dedicated to address Ofwat's customer engagement policy. This section provides clear evidence of how the research has addressed each of the points in Ofwat "PR 24 and beyond" document. The following points are satisfactorily covered: Useful and contextualised; Neutrally designed; Fit for purpose; Inclusive; Continual; Independently assured; Shared in full with others; Ethical.

Peer review

The objective of the research consists in designing, implementing and analysing a stated preference (SP) survey to estimate customers' willingness to pay (WTP) for improvements in the service provided by Wessex Water. A novel methodology is proposed to collect consumers preferences for specific attributes of interest and to measure the consumers' willingness to pay (WTP) for these attributes. The new methodology is proposed to overcome the problem experienced in the previous work where evidence suggested that respondents struggled in evaluating and choosing between scenarios defined by several different attributes and a fixed bill amount, as in the typical stated preferences settings.

The motivation to set up this new methodology is correct. Individuals have limitations in their capacity to process information, and when presented with a complex task, it is then likely that they show disengagement, adopting simplifying strategies to reduce the mental effort required to solve the problem. On the other hand, simplifying the survey tasks to reduce the cognitive burden for respondents can be seemingly perceived as unrealistic and lead to disengagement. The methodology proposed takes into consideration both aspects allowing respondents to build their preferred complex package by evaluating first each attribute alone and then re-evaluating them all together, once the package is defined. The methodology proposed is appropriate for the objective of estimating WTP for improvements in the WW services, in line with Ofwat standards.

The design of the stated preference survey is developed ensuring a strong involvement from the customers. This is in line with Ofwat standards and in line with the recommended practice

Elisabetta Cherchi

Professor of Transport, Newcastle University
Adjunct Professor, Beijing Jiaotong University, China
Editor in Chief *Transportation Research Part A: Policy and Practice* (Elsevier)
Past Chair *International Association of Travel Behaviour Research* (IATBR)

in scientific research. The initial design is defined paying particular attention to attributes that respondents consider useful from them. A series of co-developed workshops or in-depth interviews is then organised to test that the attributes and their definition are correctly understood and to understand customers opinions with regard to the attributes and the factors that affect these opinions. Clear evidence is provided that their views has been properly incorporated. A wide range of demographic characteristics are represented in this initial qualitative research, for the relevant segments of household, vulnerable household and non-household customers. Particular attention is paid to vulnerable household customers, where a specific process is designed to ensure inclusiveness. A series of post-hoc cognitive interviews were also conducted on the first draft of the survey, again ensuring inclusiveness and usefulness. Evidence is provided of how results from this step have been incorporated in the research. In addition to that, a qualitative research is conducted after the survey to better understand customers' preferences for the status quo. Reference to scientific literature is provided to support this discussion and the recommendations provided.

In line with the state of the art in the scientific community, the stated preference experiment is customised to customers' current experience. All attributes have been defined against the status quo. A methodology is defined to identify the current water bill and who is responsible for paying the bill in the household. This is critical because realism can be jeopardised by contacting respondents who are not the right persons to interview. Possible bias in the measurement of the current bill is investigated in depth and corrected.

The protocol followed in collecting the data is in line with Ofwat's recommendation. It is carefully designed to reduce possible bias mainly due to different survey formats and representativeness of the WW's customers. These effects are also controlled in the WTP estimation and research includes a transparent discussion around the potential impact of these effects and the extent to which these effects have been controlled. The response rate achieved is particularly high compared to previous similar studies (between 1.9 and 4 times higher) and the overall sample size allows for segmentation of the results based on customers' socio-economic characteristics and attitudes.

A series of qualitative techniques and other validity tests are employed to assess and correct for customers protest attitudes, status quo bias, hypothetical bias, under/over-representation of some population segments, disadvantages groups. These tests confirm the accuracy of the methodology used and the results obtained.

It is also worth noting, how all concerns raised by both the customers and relevant organisations involved in the co-creation of the survey have been considered, handled by implementing specific tests and/or analyses, and corrected or ruled out based on evidence.

Elisabetta Cherchi

Professor of Transport, Newcastle University
Adjunct Professor, Beijing Jiaotong University, China
Editor in Chief *Transportation Research Part A: Policy and Practice* (Elsevier)
Past Chair *International Association of Travel Behaviour Research* (IATBR)

2. Detailed comments and some questions

This section includes some detailed comments and questions on specific parts of the documents.

2.1 Service Attributes Selected for Evaluation

1. The definition of the attributes E and F is very generic. It has been tested that respondents understood all attributes, but it can still be difficult for the respondents to evaluate them. Attribute F also includes interventions in different areas, one is the leakage that is WW responsibility, another is customers' misuse of the water, etc. It is not possible to know if respondents evaluate all these aspects or only some of them.
2. Attributes C: this seems something that depends on how customers behave. It might not be relevant to all customers. If a customer believes that s/he does not flush wrong things in the toilet, s/he might not experience flooding.

2.3 Structure of Survey Questionnaire

3. An initial screening section was set up to ensure to record only responses from billpayers. How was the billpayer in the family identified? The actual billpayer might not be the only person who takes decision with respect to the water usage.
4. After the stated preference (SP) exercise, questions were asked to assess whether the sample was representative of the WW customer base. This means that if there is an over representation of some groups it is not possible to screen them out. For example, in section 3.2.2. it is mentioned that Metered customers are over-represented. This could be avoided if this question was asked before the SP.

2.4.1. Overview of the stated preference exercise

5. Regarding the new SP methodology proposed, it seems that there is not an experimental design, but for modelling purpose, all possible combinations are considered. Were dominant/dominated alternatives also included?
6. In the new methodology proposed, it seems that respondents do not make a trade-off between attributes. Unless respondents go above their budget, the trade-off should be only between each single attribute and the price, not between attributes. This is relevant because only 7 per cent of the sample opted to change their decision (in page 40).

3.1.1.1. Main survey

7. The size of the sample collected is good and the response rate higher than similar studies. Representativeness was tested based on socio-economic information. It is mentioned that data from APS are used to construct an age and gender profile. It is also

Elisabetta Cherchi

Professor of Transport, Newcastle University
Adjunct Professor, Beijing Jiaotong University, China
Editor in Chief *Transportation Research Part A: Policy and Practice* (Elsevier)
Past Chair *International Association of Travel Behaviour Research* (IATBR)

mentioned that data from Census 2011 are used to construct the SE profiles, but it is not specified which data exactly.

3.1.1.2. Top-up survey

8. The top-up survey was conducted in parallel with the main survey. However, the under-representation anticipated might not be confirmed once the sample is gathered. For example, later in page 36 it is mentioned that the sample suffers from under-representation of working-age C2DE individuals. This could have been corrected if the top-up survey was conducted after the main survey is completed and the sample characteristics checked.
9. Is the distribution of the socio-economic groups between the sample and the population, statistically different?

3.4. Conclusions on Survey Performance

10. In the fourth paragraph, it is mentioned that the willingness to pay analysis can accurately identify differences in customers' attitudes across demographic groups. Do you mean "customer' preferences"?

4.1. Methodological Approach

The methodological approach is clearly described, in particular it is appreciated the effort to explain it in simple terms and with a simple example. Some questions:

11. In section 4.1.4. only the incremental improvement is mentioned. How are large improvements dealt with?
12. The discussion on the negative WTP is not perfectly clear. It is explained that negative WTP can occur when respondents are more likely, on average, to choose packages with lower service levels for those attributes than packages with higher service levels, even when the total cost of the package is controlled for. In this new SP proposed, is the assumption that respondents make a trade-off still valid?

4.2.1.2. Customers prefer the status quo service level for many attributes

13. The following comment would require some clarification: "Although on aggregate our WTP results suggest that the average customer requires compensation for improvements in service, no individual survey respondent has actually expressed a desire to be paid to receive improvements in service." Does this mean that the aggregate WTP is negative but no single respondent has a negative WTP?

Elisabetta Cherchi

Professor of Transport, Newcastle University
Adjunct Professor, Beijing Jiaotong University, China
Editor in Chief *Transportation Research Part A: Policy and Practice* (Elsevier)
Past Chair *International Association of Travel Behaviour Research* (IATBR)

4.2.1.4. Findings from qualitative research on status quo preference

14. The last bullet point in page 57 suggests that a no marginal number of customers have changed choices after seeing the full budget for all the ten attributes. But in page 40 it is said that only 7 per cent of the sample opted to change their decision.

4.2.2.2. Adjusted model with controls for demographic and billing characteristics

15. Does the location of the houses have an impact on the WTP? Those living in cities might be less interested in flooding.
16. It could be (or have been) important to check if participants have experienced in their life the problems listed in Table 2.3, or have heard of people they care of or people in their area who have experienced this problem.

4.2.2.3. Simple model for sample sub-groups

17. It is found that the top-up sample respondents have lower WTP across the board than respondents in the main sample. The top-up samples has specific socio-economic characteristics. Was the WTP of the top-up sample compared with the WTP of the respondents in the main survey with the same characteristics of the top-up sample? The same for the Wessex panel.
18. The methodology used to measure differences in WTP between those who hold protest attitudes and the remaining of the sample is not clear. The protest attitude is measured using two statements evaluated on a Likert scale. If the WTP varies as a function of the protest attitudes, this should be a function of the latent variable.

5.2 Summary of Willingness to Pay Estimates

19. It is mentioned that only the results statistically significant are included in the final model. The level of significance adopted is not mentioned, but I assume this is 5% or lower. The WTP are computed and discussed with details. Many results and related comments suggest that there might be income effect (e.g. comment on the social tariff recipients, on those who struggle to pay or on the relatively “advantaged” groups). In Appendix A it is mentioned that 24% of the respondents did not answer the question about income and it is likely that these are mostly in the high range of income. Nevertheless, since the sample is quite large, even excluding this 24% it should be possible to run some tests to measure the effect of income.

Elisabetta Cherchi

