Review of the report "Willingness to Pay for Water Services at PR24 [DRAFT] Prepared for South Staffs and Cambridge Water"

by Dr. Silvia Ferrini 15 December 2022

***Overall remarks**: The report presents an interesting new valuation study to respond to Ofwat and CCW's challenge to simplify the customers' assessment of water services for PR24. The report summarises findings including multiple details of the methodological approach and occasionally it is too technical or too vague although because it is still in its draft version, I am ignoring these aspects in the review and just include specific comments. The application presents an interesting experimentation where the respondents' task is highly simplified and the overall WTP results seem sounded although further information to test their external validity might be needed. A triangulation of findings is surely a necessary step before using the results to inform developments of companies' business plan. Given the innovative design and estimation strategy and the lack of external validity, I express a circumspect opinion on the overall validity of results. This is not a criticism of what is done here but a generic problem introduced by Ofwat to welcome innovation but without specifying how to conduct a sensible validation.*

Background considerations

The water services fall into the category of non-market goods and valuing their benefits for developing a business plan is a challenging exercise that has been tested/revised and criticized for the last two decades in the water sector. Recently Ofwat and CCW have invited companies to simplify the valuation exercise following customers' feedback and regulator best practice. The report skilfully sets the scene of the valuation preliminary conditions that NERA is facing.

Strengthens of the report

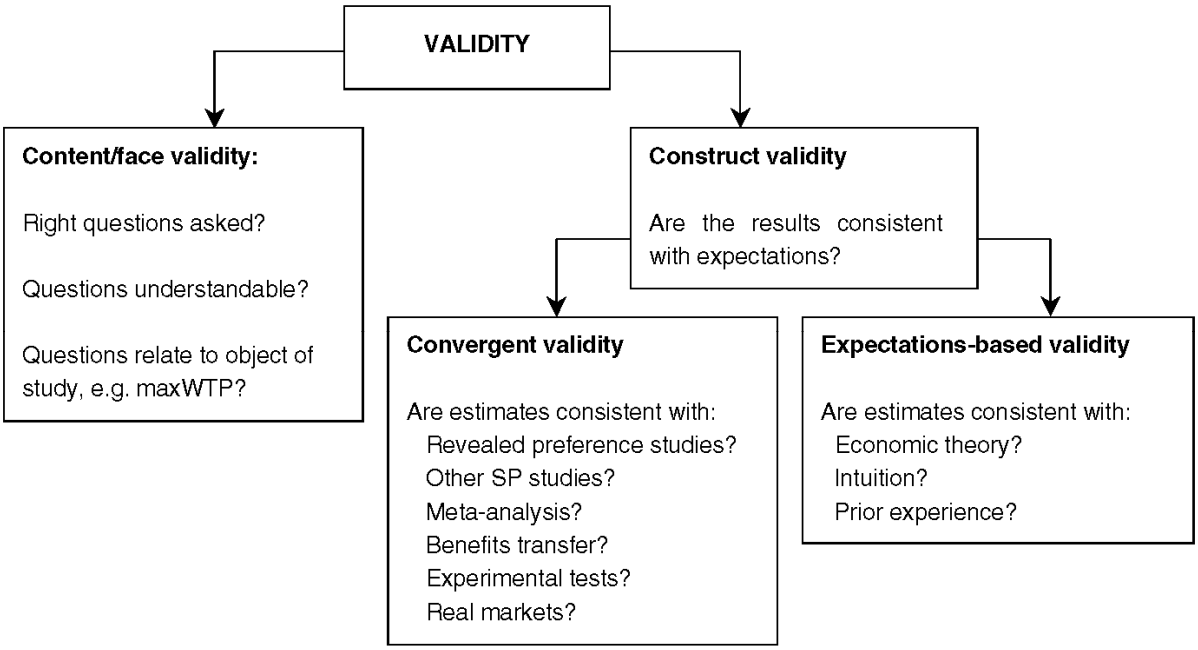
Section 2.7 provides the operative guidelines that NERA is following to develop the study and simplicity of customers' choice is their overarching approach. The study design still includes 12 water quality services that inform the business plans and respondents need to reveal their preferences for them which makes the design of the stated preference quite challenging given the pre-condition to be simple. NERA proposes a novel approach where each attribute is assessed individually according to a random design of levels anchored to respondents' current water bill. Subsequently a bundle of services with the overall price is shown and respondents can revise this final choice multiple times. Multiple interviews, testing and pilots are conducted before the main survey is launched and this provides good "internal validity". Three main categories of customers are recruited through multiple strategies for a total of over 1600 respondents and the engagement with the survey materials across the board is quite good. The table below summarizes the number of protesters, the level of understanding of attributes, the ability to compare levels and the number of people who revised the final choice multiple times. This is quite a good demonstration of the survey's ability to guide respondents in the valuation process.

	Protest (%)		Understanding (%)		Comparing attributes (%)		Revising the final choice (%)	
	SST	CAM	SST	CAM	SST	CAM	SST	CAM
HH	21	22	93	95	70	69	19	14
FBP	28	27	80	80	60	53	31	22
NHH	18	17	97	95	75	61	36	27

Two set of responses are provided by respondents:

- Individual services preference
- Overall water fee willingness to pay informed by a sequential design based on respondents' characteristics and preferences

A split choice setting is quite novel and the necessity to jointly estimate the two set of information will emerge, although in the current version the report mainly focuses on the two exercises individually. The report rarely mentions the validity of results and this is quite unusual since most stated valuation studies dedicate a significant number of pages to test validity which is generally branched in:



It is easy to claim that this survey passes the content validity since multiple assessments of the survey questionnaire have been done and the overall understanding of the study materials is very good. However, less is known about the construct validity of the study and therefore the capacity of the study to provide valid WTPs for business decision making.

Weakness of the report

The report very carefully sets out the approach for choice cards' design and estimation strategies but there is a lack of evidence that any of these novelties rely on consistent statistical and economic theories. The random allocation of levels with ad hoc rules fits the need of the survey but it is not compliant with any experimental design theory, and this might introduce systematic biases for the

final results. The same applies to the choice card setting, it is correct that realism and consequentiality reduce hypothetical biases but the possibility to select attributes individually and revise the final budget allocation multiple times violate the incentive compatibility principle (Carson et al 1997 and Carson and Grooves 2007). It is therefore unclear how this analysis can achieve expectations-based validity. The econometric model also relies on the assumption that the customers could be aware of all possible attribute combinations and select the best choice. This is again a cognitive process which has never been assessed and it is just assumed with the potential risks to introduce systematic biases. Consequently, the conditional logit model with random selection of choice set options seems problematic, since the property of the estimated parameters holds only under specific assumptions. It is reassuring that when comparing the overall model with single attribute results only a few dissimilarities emerge. Thanks for testing and comparing these two models. However, at the moment, the results appear in a sort of "black box" format since none of the models report the overall goodness of fit and WTPs' confident intervals. Finally, the report includes the analysis of the heterogeneity of preferences both with split sample analysis (4.2.4.1) and model interaction terms (4.2.4.2) and these are included in the "robustness alternative specification section". I might be mistaken but the robustness test of a model is not intendent to explore heterogeneity of preferences but to explore the stability of results to model specification strategies, therefore I would have expected an alternative functional form and/or error term specification (like a system of equations..). Finally, the WTPs results are not formally compared to any previous studies, therefore the external validity tests cannot be confirmed. Indeed, the report states that results are similar to other studies for PR24 but a formal a transparent comparison of results will be needed to reassure that estimates can be used in business decision making. For the reasons I outline above, I would suggest further independent peer review process before business decision making. It may be, for example, that my own sceptical view of implementing this interesting novel exercise to calculate WTP values is a minority view. But it is just as likely you get even more vociferous reaction in the stated preference/environmental economics/choice modelling community. It is important to garner a wider range of these views, not necessarily in terms of the current report but in terms of the validity of design and estimation setting.

References

Bateman, I., & Department of Transport Großbritannien. (2002). *Economic valuation with stated preference techniques: a manual* (Vol. 50, p. 480). Cheltenham: Edward Elgar.

Carson, R. T., Groves, T., & Machina, M. J. (1997, July). Stated preference questions: context and optimal response. In *National science foundation preference elicitation symposium, University of California, Berkeley*.

Carson, R. T., & Groves, T. (2007). Incentive and informational properties of preference questions. *Environmental and resource economics*, *37*(1), 181-210.

Rakotonarivo, O. S., Schaafsma, M., & Hockley, N. (2016). A systematic review of the reliability and validity of discrete choice experiments in valuing non-market environmental goods. *Journal of environmental management*, *183*, 98-109.