**Review of “Outcome Delivery Incentive Research: Design of Methodology” (Accent & PJM Economics)**

Professor Giles Atkinson, February 2022

**Opening comments**

The research objectives of the current project, by Accent & PJM Economics, fulfil a number of important aims and innovations in use of stated preference methods for valuing service changes in the water industry as part of the Price Review (PR) process. The ambition for a consistent survey instrument across different applications is especially important. The diversity of studies used in the past can’t have helped comparability of values across the industry either within and across PRs. The present study will address that by designing a new and single survey instrument. A whole new survey approach is possibly not the only way to arrive at a consistent set of service values. But this does have the added advantage of addressing methodological concerns that have been identified with regards to scope sensitivity and “package effects” etc. In addition, it means that updating in future can also be done in a more coherent and consistent way.

In what follows, I offer some comments on the current study report and survey under three themes: valuation methodology; selfish/ unselfish preferences; and, valuation updating.

**Valuation Methodology**

The method proposed introduces useful innovations which are distinct but related. The utility scale simplifies the choice task for respondents in a meaningful way. The certainty approach (e.g. suffering some service failure with certainty) eliminates, at a stroke, what appears to be a key source of difficulty in previous PR-related studies (numerous small risk changes). The compensation approach to eliciting customer values seems a credible way of obtaining values for attributes and levels in terms of service improvements.

My co-reviewer, Prof. Ken Willis, has already provided some excellent comments regarding the analysis of the relative impact of different service failures, and how these are to be scaled with the money values elicited for (two) service failures. I don’t have anything further to add to that, so I will focus here on other issues. But I would add my support to that of Prof. Willis: this approach seems a neat and useful way of simplifying the cognitive choice task for respondents. Doing this in a way which asks respondents to think about service failures happening *with certainty* is also welcome. Previous studies do seem to have been overly reliant on having respondents think about large arrays of attribute levels all involving small risks. It should be no surprise that this appears to have caused problems for interpreting valuation results in the past.

In terms of valuation, it is worth noting that this utility scaling could be done using willingness to pay (WTP) or willingness to accept (WTA) values. The current project chooses the latter, but it would be possible to recast all this in terms of WTP to avoid service failures with certainty. I think this chosen approach is sensible, however, for reasons that I will mention below. Clearly though it must be acknowledged that the direction of travel in much of stated preference practice has avoided eliciting WTA, and indeed it is notable that WTP seems to have been the focus of previous valuation in the water industry and price reviews in England. Therefore, it is worth Accent & PJM Economics spending some time in associated project reports providing the detailed case for using WTA.

A key reference here is the excellent review and perspective offered by Kim *et al*. (2015). This paper does not seem to feature so far in the Stage 1 report. It would be worth consulting. Broadly speaking, this paper rehearses the two main considerations governing the choice of WTA or WTP: economic principles (notably, property rights) and empirical/ practical considerations. This viewpoint is echoed in the stated preference guidelines in the influential journal article by Johnston *et al*. (2017) and in texts on (environmental) CBA such as Atkinson *et al*. (2018).

In the case of the principles for choosing WTA, in the current context, it feels like a reasonable argument that there is a property right such that water customers are entitled to not be subjected to sewer flooding, supply interruptions and so on. In that case, WTA is the appropriate valuation concept. That said, for at least one of the types of service improvements/ failures this is perhaps more ambiguous: specifically, for river water quality and low-flow (both near/ far). But, as a practical matter, it seems worth keeping in this form unless there is a strong argument for a separate valuation exercise.

In the case of practicalities of choosing WTA, these boil down to challenges in value elicitation and behavioural considerations. These are issues, for example, that might drive WTP/ WTA disparities although some of these sources of disparity should be a source of concern for valuation practitioners, but quite a number shouldn’t be (see, Kim *et al*. 2015 for a nice summary of this). For example, a concern might arise if WTA is to be elicited but the elicitation mechanism is not incentive-compatible. It’s hard to see that this sort of problem arises in the current context. But it might be worth adding some debrief question(s) on the credibility of the choice/ valuation scenarios. At present, unless I missed something, the survey does not seem to include checks on these sort of specificities.

**Selfish vs. non-selfish preferences**

The approach to be taken, in the survey, is based on asking respondents to answer based on the wellbeing of their own household. I think this is the right approach to take, given that most of the attributes of interest are very much customer-focused.

Of course, we can’t rule out that respondents will consider others as well in choosing between options. We might expect that in the case of *actual* flood incidents or water supply shortages, for example, that neighbours etc. might rally round. Moreover, there are public good characteristics of some of the dimensions being looked at here (e.g. river water quality). It might be worth thinking about adding a question, in the debrief section, about whether respondents actually thought only about their own household or not.

This information would be useful to know. But even if some respondents do indicate non-selfish preferences, their responses are likely still admissible from the perspective of value estimation/ aggregation. It depends though on what is assumed about the type of altruism from which non-selfish responses originate. That is, whether altruistic preferences imply “double-counting” really depends on what it is that a non-selfish respondent cares about. If it is specific goods that others receive (e.g. river water quality/ flooding events) then double-counting does not arise. If it is the overall wellbeing of others that people care about then it possibly does (that is, this altruism reflects concern about the costs of goods for others as well as the benefit, in terms of their provision).

These are relatively tangential/ minor matters in this context, I feel. But I mention as the report does touch on these issues and this text could do with revising for accuracy and completeness. Note also that these issues about altruistic preferences are likely to be conceptually distinct to the point made in the report about “citizen preferences” (in the sense of Sagoff etc.). Those are probably preferences best elicited within a deliberative setting (e.g. citizen juries) rather than stated preference surveys.

**Updating Values**

One last thing I wonder is the regularity of value updating that might now be required. An advantage of the method proposed is that regular/ consistent updating should be more straightforward. This aligns nicely with the practice that appears to have evolved that original valuation studies are required and undertaken in each PR period. A natural question to ask what interval is really needed for updating. Admittedly this is a question which seldom seems to be asked in environmental CBA and valuation (although see, Aldy *et al*. 2021). But, in the current context, unless we really do think values are constantly changing, we might expect a need for update only if there is substantial methodological progress or substantial time has elapsed since some update was undertaken.

There is no unequivocal answer about what this interval is, but it is perhaps worth noting that UK Government seems content to use a value of statistical life (VSL) estimated in the early 2000s. The equivalent VSL for the US dates back even further than that. Whether this has relevance for the current context is another matter. But it is worth asking whether now putting these water industry valuation studies on a more consistent footing also implies that updating might only be needed say every other price review.

**References**

Aldy, J.E., Atkinson, G. and Kotchen, M. (2021) “Environmental Benefit-Cost Analysis: A Comparative Analysis between the United States and the United Kingdom”, *Annual Review of Resource Economics*, 13: 267-288.

Atkinson, G., Braathen, N.-A., Groom, B. and Mourato, S. (2018) *Cost-Benefit Analysis and the Environment: Further Developments and Policy Uses*, Organisation for Economic Co-operation and Development (OECD), Paris.

Johnston, R.J. et al. (2017) “Contemporary Guidance for Stated Preference Studies”, *Journal of the Association of Environmental and Resource Economists*, 4(2): 319-405.

Kim, Y., Kling, C.L. and Zhao, J. (2015) “Understanding Behavioral Explanations of the WTP-WTA Divergence Through a Neoclassical Lens: Implications for Environmental Policy”, *Annual Review of Resource Economics*, 7: 169-87.