**Outcome Delivery Incentive Research: Design of Methodology.**

**Stage 1 Report**

**Report to Ofwat and CCW**

**By**

**Accent and PJM Economics**

**14th February 2022**

Reviewed by Ken Willis

Newcastle University

**Research aims and objectives**

Concerns have been raised by Ofwat [Office for Water Services (UK Government)] and CCW [Consumer Council for Water], and others, that the stated preference (SP) methodology used in previous price reviews (PRs) may be flawed, and that some results could be unreliable. This concern has arisen notwithstanding the fact that the SP willingness-to-pay (WTP) studies used in previous PRs have provided within-study evidence of their validity and reliability, as well as being subject to rigorous external scrutiny.

This concern about SP reliability has arisen mainly because valuations for specific water service measures, such as avoided internal sewage incidents, amongst others, have varied excessively across companies. Ofwat and CCW propose ensuring comparability of results across water companies by applying a common methodology, and thereby identifying genuine differences. In pursuance of this, the brief for the proposed research emphasises that the methodology should focus on consumers, consider new research ideas, inform output delivery incentives (ODIs) and performance commitments (PCs), by valid and reliable, and comparable across water companies.

**Research Methodology Review**

To fulfil these aims and objectives, Accent and PJM Economics have proposed a methodology grounded on impact based evidence, to estimate the relative impact of all types of service failure; coupled with a compensation choice approach, to estimate the values for different types of avoided service failure.

Section 3 of the Report, on “Survey Design”, provides an excellent account of the advantages and disadvantages of the different SP options, namely discrete choice experiments (DCEs) plus contingent valuation CV); best-worst scaling or maxdiff methods, compensation-based valuation, and menu or slider choice methods, in the context of research for PR24. The Section also clearly outlines the selected SP approach; issues in the definition of service measures; choice formats; experimental design; and questionnaire structure.

There is considerable merit in the compensation-based approach proposed by PJM Economics and Accent. It is simple for consumers to understand; and avoids the need to specify numerical levels for each service measure as in previous PR SP studies based upon a WTP methdology. Customers do not have to judge their WTP for a specific numerical improvement in a service level improvement (e.g. 450 properties rather than 540 properties not subject to internal sewage flooding out of say 3 million customers). Such a service improvement which would, in all probability, not affect them during the ensuing PR period.

Previous SP studies often expressed service changes in probability terms (either low probability high consequence e.g. internal sewage flooding; or higher probability but low consequence e.g. a water supply interruption for 3 hours). Very small changes in risk (expressed as *X* per 100,000, or *X* per 10,000, etc.) are difficult to compare, judge and value. It is much easier for consumers to judge between a pair of service failures, which service failure would impact them most, and hence which of the two they would prefer to avoid. It is also probably easier for consumers to judge the amount of compensation required for an event or service failure, which has impacted them, to restore the consumer to his/her original utility level.

A compensation-based methodology will probably result in less variation in marginal benefit values for service measures across water companies. It will work with many water service measures such water supply interruptions, consumer contacts about water quality, internal and external sewage flooding, bathing water quality, and river water quality.

A problem with the approach might occur for some factors e.g. discharge compliance, and environmental performance assessment, where these have an indirect impact on customers; while pollution incidents are probably confounded with bathing water quality and river water quality (as noted on page 50 of the Report). Compensation for factors such as these might be more difficult for consumers to judge. But the application of an impact-based exercise to estimate the relative impact of all types of service failure, in relation to the compensation value for one water supply service measure [unexpected water supply interruption for up to 6 hours], and one waste water service measure [sewer flooding outside the property (house) but within the property boundary], [page 51 and 52of the Report], will ensure values for such services.

It is proposed that compensation values are derived from for one waste water service measure and one water supply service measure for the remaining service measures. This will probably result in two different benefit values for the same service measure, and, as the Report notes, would be especially likely to occur if one was linked to compensation for internal sewage flooding, whilst the other is linked to say a planned water supply interruption for 6 hours. The “impactfulness ratios” would need to reflect the difference between the compensation required for the two service factor failures selected if similar ball-park values are to be derived. Much will depend on the calibration of the compensation values displayed on the choice cards. Using “unexpected supply interruption”, and “external sewer flooding incident” (as suggested on page 51 of the Report) should ensure greater convergence of benefit values from the two respective compensation measures, than using internal sewerage flooding.

Several SP studies for PR14 and PR19 noted that consumers were unwilling to trade-off decrements to service for a bill reduction. This implies losses (decrements to service) are valued greater than gains (improvements) to service levels. This is predicted by prospect theory (Kahneman and Tversky 1979), encompassing a neutral reference point, diminishing marginal utility, and loss aversion. Thus, WTA compensation values for a service decrement may over-estimate the value of WTP for improvements to households for water service measures. This is more likely to occur when values are derived for risks of incidents occurring, especially when risks are very small. People have difficulty in judging, and valuing, very small risks (see Jones-Lee et al, 1985). However, the methodology proposed is not based on risks, so there is less chance of such an over-estimate in value occurring.

The advantage of the compensation approach is that, as the Report points out (page 58), it concentrates on the utility difference between experiencing and not experiencing a service issue, rather than via the utility difference between a small risk of experiencing a service issue and a slightly smaller risk of experiencing the same service issue, which characterised benefit estimation in previous PR studies. Individuals commonly make errors in calculating risks (probabilities and expected values), when thinking fast (as in responding to a SP questionnaire), compared to thinking slowly and calculating the difference. Hence the compensation approach has considerable merit in seeking to move away from a method based on individuals’ WTP to experience a smaller risk of a water service failure.

On the other hand, a compensation approach may result in a lower gross benefit estimate for some service measures than a WTP approach, since values should be aggregated across consumers who have been impacted, and across the number of consumers who are estimated to be impacted over the quinquennial period. In comparison the WTP approach aggregates benefits, albeit a smaller estimate of benefit reflecting the low risk of occurrence for most service measures, across all water company customers. Both the WTP and the proposed WTA approach seek to exclude non-paternalistic altruism, but both can include option value and non-use value for service measures, depending on how the questions are framed.

The Report (page 52) stresses that households, in judging values and impacts, should consider any concerns they may have for the local area or environment; and ignore impacts on others outside their household. I agree entirely with this approach: altruism is easily double counted, and this needs to be avoided in any valuation exercise.

The Report discusses the experimental design for the study. The experimental design, and the analysis of the data, needs to ensure that the compensation values, based on impactfulness coefficients derived from the impact based exercise design, do not simply reflect scaling in the data. The Report addresses the scaling issue (page 72), but water companies need to exercise caution interpreting service measure utility coefficients from the models in terms of making comparisons between different water companies. Because of scale effects, in what will effectively be different surveys across different water companies, the utility coefficients may not be directly comparable, although the WTA money values for these service measures will be.

The Report (page 77) demonstrates how value per minute can be calculated for an avoided 6-hour unexpected supply interruption. The value per minute, aggregated over the whole company, would be 100\*N/360 where N is the number of customers in total. [The avoided supply interruption was hypothesised to be valued at £100]. But compensation should only be applied to those customers who are affected by the supply interruption. As a principle, compensation is not paid to customers who are not affected by an incident. Compensation is paid, of course, in some cases where there is a small probability of harm occurring e.g. wage premiums in risky occupations. But the compensation approach in PR24 is designed to measure the impactfulness and value of incidents if a customer actually experienced it; and not for the value of a small chance of the incident occurring. However, the values can be expressed for output performance measures, over the whole customer base.

The questionnaire suggests eliciting compensation for two service measures: an “unexpected supply interruption”, and an “external sewer flooding incident”; and deriving values for other service failures by the impactfulness of other service failures in relation to these two. In a WTP approach this would result in an over-estimate of the aggregate benefit across all the service measures because of the “package effect”. The compensation approach avoids this additivity phenomenon, unless there are substitution and correlation effects between some service failures e.g. between pollution incidents, discharge compliance, storm overflows, bathing water quality, and river water quality, and possibly interaction effects between internal sewer flooding and external sewer flooding.

However, the additivity issue is unlikely to be a significant limitation in the case of water and wastewater services valuation research. Even using the WTP discrete choice experiment method in previous price reviews, which does allow the estimation of interaction effects, it is extremely rare for studies to model these effects and take them into account when making recommendations for appraisal values. In the vast majority of cases, only the main effects of attributes are modelled and valued.

The questionnaire is well constructed. I wonder if it is worth asking customers in Q14 to identify which service failures in the list the household has experienced, rather than just whether they have experience any of them over the last 3 years. This would provide a richer data set.

Also the wording in Q29 stresses the choice for the water company of spending more to reduce the number of service problems, or paying out more compensation to customers that experience problems. The wording infers that the respondent, by taking compensation, deprives the water company of investment to help other customers affected by service failures. Of course minimum compensation required can easily be to over-estimated. But the wording needs to ensure respondents do not under-estimate compensation required. The question wording also implies a concern for others [“striking a balance between the levels of compensation paid and the number of service problems experienced by customers”] which seems contrary to earlier advice (Report page 52) of ignoring impacts on others outside their household.

**Conclusions**

PJM Economics and Accent have, with creative thinking, been able to propose a new methodology for PR24. The proposed methodology is a radical departure from previous price review WTP studies, but it is still customer focused. The proposed methodology is based on WTA compensation for a service failure. This eliminates the need for customers to values small changes in risk as in previous price review WTP surveys, and instead bases values on a customers’ comparisons between different types of water supply and quality, and waste water, events. The proposed methodology should, with careful survey application, provide reliable and robust values for marginal changes in service measures. Ofwat and CCW can have confidence that the results based on the proposed methodology will provide a truer reflection of customers’ values for different service measures, with less variability across water companies.

**References**

Jones-Lee, M.W., Hammerton M, and Philips P.R. (1985). The value of safety: results of a national sample survey. *The Economic Journal* 95: 49-72.

Kahneman D. and Tversky A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*  47: 263-291.