Research on customers’ willingness-to-pay for service changes in UK water company price reviews 1994–2019

Ken Willis\textsuperscript{a} and Rob Sheldon\textsuperscript{b}

\textsuperscript{a}School of Architecture, Planning and Landscape, University of Newcastle upon Tyne, Newcastle upon Tyne, UK; \textsuperscript{b}Accent, London, UK

\textbf{ABSTRACT}

Water companies are regional monopolies in the UK, and subject to quinquennial price reviews to ensure customers receive value for money. This paper documents the application and evolution of stated preference methodology in the quinquennial price review research. Stated preference methods are used to assess customers’ preferences for changes to water supply, waste-water, and environmental services; and customers’ willingness-to-pay, or willingness-to-accept, water bill amounts for changes to these service levels. Recently revealed preference methods have been given more prominence in estimating values for some water services. The application of stated preference, and revealed preference, has seen continued attempts, in successive price reviews, to improve the accuracy and reliability of values for water services, as an input into the cost–benefit analysis of water projects, and water company business plans.

\textbf{Introduction}

Water and sewerage companies are regional monopolies in the UK, in terms of domestic customers. In England and Wales water and sewages services are supplied by private companies. Following privatisation in 1989, the first review of water company prices took place in 1994. This was the first opportunity Ofwat (Water Services Regulation Authority), the economic regulator for the water sector in England and Wales, had to set price limits. Water companies could choose to accept Ofwat’s price limits, or refer Ofwat’s decision to the Monopolies and Mergers Commission (the forerunner of the present Competition and Markets Authority).

Ofwat sets standards to which water companies must conform (i.e. customers must have an adequate water supply and water pressure, and homes must not be flooded from sewers); and also every five years it sets prices that water and sewerage companies can charge their customers over the ensuing quinquennium. Ofwat sets prices to enable water companies to make a reasonable return in their investment: around 6% per year. Failure of companies to meet guaranteed standards can result in Ofwat imposing fines on companies, and ultimately to the withdrawal of their licence to operate as a water and sewerage company.

Water companies also have to conform to legal minimum standards set by the UK government (e.g. by the UK Drinking Water Inspectorate, responsible for regulating the quality of drinking water; and the Environment Agency for protecting and improving the quality of rivers, estuaries and coastal waters). Water and sewage companies also have to conform to European Union
(EU) Directives (e.g. the EU Water Framework Directive (WFD)(Directive 2000/60/EC) required all inland and coastal waters to reach ‘good status’ by 2015, constraining the chemical and biological content sewerage companies can discharge into rivers and coastal waters). However, water companies have scope to vary service levels provided beyond legal minimum standards to satisfy customer demand, where benefits to customers can be clearly demonstrated to exceed costs. Many water companies exceed minimum legal standards on some water supply and water quality service factors, but such additional investment providing higher standards needs to be justified to the regulator, Ofwat.

Water service and quality service measures include interruptions to supply (planned and unplanned), security of supply, discolouration, water taste and odour, sewage flooding (inside and outside the property), odour from sewage treatment works, river water quality, and bathing water quality, and pollution incidents (accidental discharges from sewage treatment works). These services are essentially ‘public goods’. For many of these services individual customers cannot purchase different levels of service: everyone on the network receives the same level of service, and everyone living in the water company area faces the same environmental levels (river water quality, and bathing water quality), except for local variation in distance from the customer’s property.

Public goods, or non-market goods, can be valued either by stated preference (SP), revealed preference (RP), or preventative expenditure and avertive behaviour costs. All have their strengths and weaknesses. HM Treasury (2020) provides examples of the applicability of these methods in valuing different environmental and amenity goods, including RP methods to value flood prevention.

Some of these water service attributes, such as smell from sewerage treatment works (STW) and river water quality, can be valued by revealed preference (RP) methods, namely changes in house prices, or by purchases of bottled water to avoid taste and odour of water companies’ drinking water. But other attributes, such as interruptions to supply, and security of supply, can only be measured by stated preference (SP) non-market valuation methods. Indeed, it was often easier and more practical to value all the service measures using SP rather than have separate RP and SP studies for different attributes. Moreover, market measures, e.g. purchases of bottled water, do not capture all the utility loss from taste and odour of drinking water, failing to include consumer surplus on existing purchases, and ignoring utility of consumers who would pay an amount greater than zero for bottled water but not the existing market price.

The valuation of changes in service provision, can be set against costs, allowing a cost–benefit analysis (CBA) of all potential investment projects, over the price review (PR) period, to be included in the water company’s business plan. Such a CBA, enables water companies to maximise net benefits to customers and shareholders. Economic optimisation of all investment projects within the company, which may cover thousands of projects over a 5-year period, involves ranking projects in terms of benefits relative to costs. Optimisation also involves water companies producing a cost schedule for all investment projects, including the incorporation of risk of asset failure (e.g. probability of a mains water burst at various points along a network), to compare against the benefits of undertaking maintenance or improvement projects.

The purpose of this paper is to document the history and evolution of the application of methods, mainly non-market stated preference (SP) methods, to estimate customers’ willingness-to-pay (WTP) for changes in water companies’ water service standards, as an input into financial and cost–benefit analysis (CBA) appraisals of projects for each quinquennial investment period, and price review by Ofwat. The extensive use of non-market valuation in the water industry in the UK has led to many different applications of SP methods. These applications have drawn upon the existing theoretical development and uses of SP in the literature, rather than specifically being theoretically innovative in their own right. In addition, the detailed application of SP has varied from one price review to another, partly to address the issues encountered in the application of SP in the previous price review; and partly as a reflection of the evolution of SP methods over time since the 1990s. The paper therefore sets out in chronological order through each price review (PR), the application of non-market methods to values service improvements.
Price review 1999

The first use of SP (stated preference) analysis in water company customer research, to assess customer demand for improved service levels over and above minimum legal requirements, was that undertaken by Accent for Yorkshire Water’s PR99 (price review 1999). The PR99 determined the asset management programme (AMP3), and business plans, for all water companies, over the next 5 years (2000–2004). In determining the quinquennial price charged to customers, water and sewerage companies have to submit a business plan to Ofwat. This includes information on investment required to meet legal standards, productivity gains in service delivery, and information on customer preferences and demand to justify any investment beyond statutory requirements.

The stated preference (SP) approach had, prior to PR99 mainly been used in transport, plus a few environmental economics cases, but became a key pillar for the quinquennial water price review research programme. The SP study for Yorkshire Water for PR99, which was fairly rudimentary, used a standard conditional logit (CL) model, but it identified many of the issues which later dominated subsequent price reviews: estimating willingness to pay (WTP) for better service amongst some water services, but not all; grappling with respondents’ abilities to trade-off quite different and complex issues; the need to scale the WTP values; the necessity of encompassing both household and non-household customers; a concern about environmental matters; and respondents’ ability to trade-off temporal issues.

Concurrently with PR99 and AMP3, the EU (European Union) Water Framework Directive 2000/60/EC committed European Union member states to achieve good status for all water bodies (including marine waters up to one nautical mile from shore) by 2015. So alongside water companies’ priorities for customer service improvements for water supply, water quality, and waste water disposal in PR99, the UK Environment Agency (EA) assessed customers’ values and priorities for improvements in water bodies (rivers, lakes, and bathing waters).

In pursuit of this the NRA (National Rivers Authority) (a precursor to the EA), had commissioned the Foundation for Water Research (FWR) to produce a manual detailing how to value the benefits of river water quality improvements, based on environmental economics studies. The manual provided examples of how to estimate a variety of benefits (fishing and angling benefits, other recreation benefits, amenity benefits, etc.) associated with improvements to water quality and flow, reductions in the risk of flooding, and so on (FWR 1994).

After the 1994 price review, the FWR (1994) manual was deemed to be too detailed and cumbersome to operate by EA staff with little knowledge of economics, and hence it was abandoned. A multi-attribute technique (MAT) was then adopted by the EA to appraise schemes for PR99 (Palmer 1998). This MAT was a simple subjective points scoring method: points were determined by the degree of water quality improvement and the length of river improved.

Although easy to use by non-economists, MAT was severely criticised by economists since it is not based on economic theory, and therefore leads to inaccurate benefit estimates. The demands for a more economic approach to assessing and valuing the benefits of environmental improvements, in relation to costs, in the 2004 price review (PR04) led the EA (2003a, 2003b) to develop its Benefit Assessment Guidance (BAG).

PR04 customer WTP research

Reviews of the previous regulatory round in the water sector in 1999 identified the regulation of investment to maintain assets as a key area requiring development in advance of the then periodic price review in 2004 (PR04). Indeed, Ofwat (2000) published a letter to water company managing directors (MD161) requiring companies to develop a better understanding of the economic case behind capital maintenance investment.

In response to this, Yorkshire Water, a leading UK water and sewage company, developed an integrated decision process with the acronym LEADA (Leading Edge Asset Decision Assessment).
LEADA applied economic optimisation to all investment projects within the company (typically 20,000 projects over a 5-year period) to rank them in terms of benefits relative to costs. This involved Yorkshire Water producing a cost schedule for all its investment projects, including the incorporation of risk of asset failure (e.g. probability of a mains water burst at various points along a network) to compare against the benefits of undertaking maintenance or improvement projects. Benefits were assessed in terms of customers’ willingness-to-pay (WTP) for improved service performance. UK Water Industry Research (UKWIR), the producer organisation representing all water and sewage companies in the UK, stated in 2003 in its ‘Capital Maintenance Planning Manual: Current Methods and Good Practice Guidance’, that:

The LEADA system … should become the benchmark development of capital maintenance planning approaches for those companies or water and sewage providers which elect to follow the cost-benefit planning objective.

The Price Review 2004 (PR04) was a watershed in the application of stated preference (SP) methods to assess customers’ values for water service changes in the UK. Yorkshire Water embraced SP methodology and put a huge amount of resources into the SP research and the subsequent linear and non-linear programming procedures to maximise the difference between benefits and costs of a large number of possible improvement and asset management projects, to optimise investment and maximise benefits to customers. This research for Yorkshire Water’s PR04 won the Operational Research (OR) Society’s President’s Medal for the best application of OR in 2003 (Willis, Scarpa, and Acutt 2005). This set the standard for water companies to emulate in future price reviews and business plans.

The stated preference (SP) research for Yorkshire Water for PR04 was innovative in a number of ways. It applied more advanced discrete choice experiment (DCE) techniques in the SP research in addition to standard conditional logit (CL) models. These included CL, CL quadratic, nested logit (NL), NL quadratic, and random parameter logit (RPL) or mixed logit (MXL) models (see Willis, Scarpa, and Acutt 2005) to assess which type of model best fitted the data. Concurrently, Hensher, Shore, and Train (2005) were using MXL models to establish customers’ willingness to pay to avoid interruptions in water service and overflows of waste water, in Canberra, Australia, differentiated by the frequency, timing, and duration of these events. WTP for different water services is an important input into the regulatory process for establishing service levels and tariffs, as well as useful information for water companies to enable them to deliver services at prices customers deem ‘value for money’.

Subsequent analysis of the Yorkshire Water data explored systematic differences in preferences for non-status quo alternatives. Preferences for change versus status quo were explored with an alternative specific constant (ASC), nested logit and an error component (EC) specification. Alternatives offering changes from the status quo did not share the same preference structure as status quo alternatives. Evidence suggested that bias in estimates ignoring the status quo effect is substantial (Scarpa, Ferrini, and Willis 2005). The status quo (SQ) issue was further explored by applying an EC discrete choice model which again showed that alternatives offering changes from the SQ did not have the same preference structure as the SQ alternative, but that estimates of spread parameters in zero-mean error components can be decomposed conditional on water company customers’ socio-economic characteristics (Scarpa, Willis, and Acutt 2007).

One of the main issues which the PR04 study for Yorkshire Water had to address, was the large number of water service attributes. Yorkshire Water required estimates of customers’ preferences, utility, and willingness-to-pay (WTP) for fourteen service measures (see Willis, Scarpa, and Acutt 2005). It is not possible for customers to cognitively trade-off 14 service measures simultaneously, without adopting some simplifying heuristic, which may give rise to biased results. Studies have also suggested that choice complexity increases variance and reduces choice consistency. Consumers faced with increasing choice complexity might be less able to make ‘accurate’ choices (Swait and Adamowicz 1999); and thus their preferences might be characterised by different levels.
of variance depending on the complexity of the task. Complexity of choice may also be related to the propensity to ‘avoid’ choice by deferring choice or choosing the status quo. DeShazo and Fermo (2002) noted an increase in the variance of the error component in utility with increasing complexity of choice. An increase in the number of alternatives was found to initially decrease error variance before error variance increased with the number of alternatives.

So the Yorkshire Water SP discrete choice experiment (DCE) study presented the 14 attributes in 5 blocks of attributes, each with a money bill amount attached. Thus each block of attributes encompassed three to five attributes. Including a price in every block of attributes itself creates problems. Respondents might happily be willing to pay the money amounts in the initial blocks of attributes, then reject improvements in service attributes in later blocks presented to them, as they realise the totality of the bill increase across all blocks. So the order in which blocks are presented to respondents can affect customers’ utility weight and WTP values for the attributes. It is not possible to link utility weights across blocks, hence the rationale for including price in each block. This gave rise to another problem in DCE, which has plagued contingent valuation (CV) studies valuing goods: the individual valuation and summation (IVS) issue.

Where there are many attributes, more than one CE is often used to value a good. This can give rise to the independent valuation and summation (IVS) problem (see Hoehn and Randall 1989). Various studies have shown, both in theory and practice, that an IVS approach grossly overestimates aggregate WTP when valuing goods independently where there are significant substitution effects (Hoehn and Loomis 1993). Additive separability1 is a common assumption in aggregating the values for different goods in space and time. But where there are substitution effects between goods then additive separability, or IVS, cannot be applied.

The IVS issue was addressed in PR09 studies, where researchers typically included a ‘package’ value DCE (with all attributes included), and/or a contingent valuation (CV) question, to estimate customers WTP for the total package of improvements across all attributes. The package DCE was often a simplified DCE (with a limited number of attribute levels: typically only three levels, the status quo, +1, and +2 levels, to reflect the current situation, a practical service improvement level, and a more aspirational improvement level) to assess WTP for the total package of attributes at level +1 and +2. The package effect was seen in the PR09 study by Burge, Tsang, and Rohr (2007) for Northumbrian Water, where the package values for each service were only 26–37% of the values for each service attribute derived from the three DCE lower level DCEs; and only 24–56% of the lower level DCE service values for Essex and Suffolk Water.

Meanwhile the Environment Agency’s (EA) BAG manual was criticised following PR04. Most PR04 environment schemes were statutory schemes requiring water bodies to meet EU Water Framework Directive standards. The benefits assessment guidance (BAG) was applied to all schemes where there were choices to be made (CTBM). These were schemes where improvements could be made above statutory minimum standards, if benefits exceeded costs. Benefits included improved fishing, recreation, and natural habitats, better water quality and flows, and a reduction in the incidence of disease from cleaner bathing waters. The benefit/cost ratio (B/C) for a CTBM scheme to be justified under BAG was 1.2.

The costs for each CTBM scheme were estimated by Ofwat, from the water companies’ proposed investments and costs in its asset management plan for 2004–2009 (AMP4) (i.e. their business plans). The EA calculated benefit/cost (B/C) ratios for each CTBM based on its BAG benefit estimates and Ofwat cost estimates. The EA then ranked all schemes based on these B/C ratios. Only at this stage of the cost–benefit analysis (CBA) were respective water companies invited to comment on the economic appraisal of the CTBM schemes. BAG benefit values are based upon three principles: benefit transfer (BT), distance decay of the proportion of the population having a positive non-use value, and the independent valuation and summation (IVS) of scheme benefits.

BT is not a reliable method to establish accurate values for environmental goods. Hanley, Wright, and Alvarez-Farizo (2006) rejected BT between the River Clyde and the River Wear. They found people living near the River Clyde valued improvements to their local river more highly
than people in Durham valued identical improvements to their local River Wear; despite the fact that the Clyde survey sample had a lower income than the Durham survey sample.

Moreover BAG overestimated benefits through its reliance on IVS. The actual non-use benefit of water quality improvements for a single river were much lower than the estimated benefit derived from using BAG, owing to a failure to account for the effect of the total number of schemes in the surrounding area on households’ budgets. In addition, the valuation studies used by the EA in its analysis were those for water quality improvements in water bodies with poor water quality. Declining marginal utility would suggest the marginal value for improvements to water of already good quality would be much lower (Willis 2008).

The discrete choice experiment (DCE) undertaken for Yorkshire Water’s PR04, was used to appraise all the CTBM schemes proposed by the EA in the Yorkshire Water area. There were considerable discrepancies in B/C ratios, between those estimated under the EA’s BAG approach and those estimated through the DCE model, for schemes deemed to be ‘beneficial’ by the EA using BAG (see Willis 2008). B/C ratio estimates differed between the two methods often by one or even two orders of magnitude! This analysis again supports the hypothesis that the EA’s BAG approach grossly overestimated the value of benefits in the analysis of CTBM schemes, because it failed to address the IVS issue.

**PR09 customer WTP research**

PR09 saw discrete choice experiments (DCEs) adopted by almost all water companies. There was little or no direction from Ofwat on how a DCE was to be applied. Hence water companies adopted quite different approaches in terms of the definition of individual water service measures (e.g. length of supply interruption); how water service measures were expressed to respondents (e.g. actual number of properties flooded, percentage of properties flooded, and severity of flooding); the sample size of the survey, method of sampling (personal interview, telephone interview, mail shot, or on-line self completion). Nor was there any direction from Ofwat on which type of DCE model (conditional logit (CL), mixed logit (MXL), nested logit (NL), error component (EC), etc.) should be used. Hence there was considerable variation in the framing of the attributes, the experimental design, the SP approach, and in the DCE data analysis amongst water companies in PR09. The result of this lack of standardisation in assessing customer preferences, was a large variation in WTP values for different water service measures between different water companies. This led Ofwat to commission NERA (National Economic Research Associates) and Accent to advise on carrying out WTP surveys, to inject some standards and standardisation into future water company WTP surveys.

The report by NERA/Accent (2011) developed a common valuation framework for carrying out WTP surveys, by setting out practical steps companies can use for: (i) the identification of priority service measures, (ii) attribute and level definition and presentation, (iii) deciding elicitation methods and wider aspects of survey design (e.g. sampling, warm-up and follow-up questions to be included in the survey). This report looked at particular issues arising in PR09. These included geographic variability in service levels and uneven service improvement impacts; altruism as a driver of some valuation responses, especially for certain service measures (e.g. sewer flooding, persistent low pressure, odour from sewage treatment works); respondents’ valuation focus (e.g. whether the respondent’s property was affected or not); presentation of risk (e.g. respondents’ understanding of low probability events); and the presentation of bill and bill changes (e.g. respondents’ understanding of what bills will be after the review period).

The NERA/Accent (2011) report made a series of recommendations based on:

1. principles of how valuations are best derived (e.g. market based estimates are preferred to revealed preference (RP) which in turn is preferred to stated preference (SP) methods); relevant estimates from existing literature should be compared with the results of the WTP exercise;
(2) principles for defining the focus of the evaluation (attributes should focus on effects that are local to consumers; for improvements that most respondents in a company-wide random sample will not experience, direct estimation of the values held by those affected should be adopted, with market or RP methods; non-use values should be elicited for environmental values held on a nationwide basis; and altruistic values should also be sought);

(3) principles for defining and presenting attributes and levels (units of measure used to describe attributes and levels should be easy for asset planners to use, easy for customers to relate to, described from the respondents’ point of view, with a level of numeracy they can understand, with visual stimuli (charts or pictures) as well as text; attributes levels should involve a deterioration in service so that the value of maintaining service can be ascertained, and asymmetry of values around status quo can be tested; the bill attribute should be clear about inflation and other reasons why bills might increase and to what extent);

(4) principles for determining elicitation formats (e.g. apply ‘lower level’ and package choice experiments, coupled with contingent valuation (CV) question(s) to allow testing and scaling for packaging effects): restrict the number of attributes per choice experiment exercise to 6 or less, and restrict the number of alternative choice sets to 2–4, to prevent cognitive overload; restrict the number choice sets within any one exercise to a maximum of 8, with randomisation of the ordering of CE exercises, and restrict the total number of exercises presented to any one respondent to 4 or less, with an average interview length to 30 minutes or less. WTP surveys were thought best administered under interviewer guidance (a face-to-face or telephone interview approach). Respondents should also be reminded of all of their budget. Minimum samples of 300–400 interviews per segment of interest were recommended, with random quota sampling to ensure a representative sample of customers. Data on respondents’ service experience, usage and attitudes should be used as a reliability check on the WTP estimates; and validity testing should include follow-up questions to probe why respondents chose the way they did in the valuation task.

Intrinsic within this was an appreciation that there needed to be increased thought given to ‘market research’ issues alongside the quite substantive development of econometric perspectives. The report also advocated omitting an explicit status quo (SQ) option on the choice cards in the DCEs, forcing respondents to choose between two hypothetical alternatives. This was argued to be a mechanism for maximising information on trade-offs between attributes, by discouraging people from opting out of choices. This recommendation was controversial and not all water companies followed this practice in PR14 research.

The NERA/Accent Report also identified a subset of service measures that cannot reliably be included in WTP surveys, or that are potentially better approached with other methods. The NERA/Accent (2011) report suggested that (1) health effects of drinking water quality (excluding lead), (2) lead content in tap water, (3) sludge disposal, (4) construction impacts, (5) supply pipe liability, (6) greenhouse gas (GHG) emissions, and (7) specific habitats, should all be valued using RP or value transfer rather than through a SP approach. Some companies followed this advice and used epidemiological studies to value lead reduction in drinking water; revealed preference (RP) methods to value safe drinking water quality; RP for injuries to staff; RP for bathing water quality; and traded and social price carbon values for green house gas (GHG) emissions.

The water industry also sought to draw lessons from PR09, and commissioned Cascade Consulting, eftec, and ICS Consulting (2010) to investigate the application of CBA and benefit valuation. The Report revealed that CBA had very little influence at the policy level in terms of water company decision making. At the programme and project level, the use of CBA was determined by whether companies used CBA for strategic planning purposes or tactical purposes. In the former, CBA was used to guide choices as to the service outcomes companies wished to pursue, and in the latter CBA was used to demonstrate or justify the case for proposed investment. PR09 evidence suggested the latter was more predominant than the former. In many cases industry planners exercised caution,
even reluctance, to let CBA over-ride engineering and technical judgements about investment needs. As a result, CBA was used less to determine the outputs to be delivered by investment programmes and more to justify the expenditure associated with investment programmes.

In terms of the WTP values derived by the PR09 studies, the Cascade Consulting, eftec, and ICS Consulting (2010) Report noted that understanding of SP methods remains low in the water industry, with ‘ownership’ of the valuation work often left with the consultants. Evidence from PR09 showed the wording used in the definition of service attributes, and the selection of the units of measure of change for each attribute, were crucial factors that influenced valuations.

One worrying aspect of PR09 results was the great divergence of values across water companies: the highest WTP value for both leakage for example was 100 times greater than the lowest value. The same difference in magnitude was noted also for water quality, whilst there was a thirty times difference between the highest and lowest value for sewer flooding. The variability in WTP between studies (of different water companies) was much greater than the variability in WTP within a study (e.g. where WTP values from more than one model (CL, MXL, NL, etc.), had been reported).

A meta-analysis, of different water company studies, sought to assess the impact of study characteristics, modelling approach, socio-economic characteristics of customers, and water company characteristics, on WTP results for interruptions to supply (ITS), leakage, security of supply (SOS), water quality (WQ), and sewer flooding (SF). The analysis found that differences in study design, and methodology, contributed to differences in WTP estimates. Also the way in which attributes or service measures were defined influenced WTP values (Cascade Consulting, eftec, and ICS Consulting 2010).

Detailed research on the use of value or benefit transfer (BT) was developed by eftec (2009) for Defra (Department for Environment, Food and Rural Affairs), with an example and application to river water quality in West Yorkshire. However, it was clear from the Cascade Consulting, eftec, and ICS Consulting (2010) research that benefit transfer (BT) could not simply be taken from the individual water company SP studies for PR09, to estimate customers’ WTP in other water company areas. The dispersion of WTP values, and the inability of models to explain the WTP values, were just too great to render BT practical.

This led to the encouragement of greater standardisation of attribute measures and methodological approaches in PR14, in the hope that there would be less divergence between water company WTP estimates for each service measure. The great divergence of WTP for service measures in PR09, and the inability to explain these, also indicated that it was potentially inappropriate to institute one national SP survey, the results of which would be applicable to all water companies.

PR09 also raised questions of whether revealed preference (RP) could be employed to provide more robust or collaborative evidence on values for some water services. This led Ofwat to commission Cascade and eftec (2011) to investigate the use of revealed customer behaviour in future price reviews. The report explored how revealed behaviour (avertive expenditure, travel cost (TC), and hedonic price (HP) methods) could be used to value drinking water aesthetics (taste, odour, and appearance), health risks from drinking water, (through the purchase of bottled water), water hardness (purchase of water softeners), interruptions to supply, low pressure, odour from sewage treatment works (through differences in house prices), sewer flooding (through damage to building fabrics and contents via market prices (damage costs)); river water quality, and bathing water quality (via TC models of recreation demand), across business and domestic customers. However whilst the cost of sewer flooding could be valued through HP house values, HP will not provide reliable WTP estimates if the sewer flooding effect is not perceived in the housing market purchases.

**PR14 consumer research**

There was a large advance in the development of the PR14 WTP for water services prioritisation research, following this close evaluation of the PR09 work. In particular there was considerable debate and consideration given to a number of issues, including the use of qualitative work in
helping design the tools to be respondent friendly; the use of visual material to denote different levels of risk; whether there was a need to always include a SQ option on a choice card; the treatment of efficiency changes in option development; the treatment of inflationary price impacts over the quinquennium (whether to just express the water bill price ‘now’ i.e. 2013/2014; or the water bill price with inflation in 2019); alternative ways of eliciting scaling factors without necessarily including price in all CEs, in what became known as the lower level exercises, or whether these could be derived through an additional so-called package exercise (valuing different combinations of +1 packages of improvements, against different combinations of packages of +2 improvements, against the SQ); and how to efficiently explore a number of levels of change without needing to undertake Stage 2 surveys of severity of impacts (i.e. not only in terms of the number of people affected (e.g. by internal sewerage flooding), but also by the severity of the impacts (e.g. damp patch in cellar, standing water and sewerage in cellar and restricted toilet use, sewerage under floor boards, to standing sewerage water in living accommodation)). It was hoped that these advancements would allow respondents to make more informed and rational choices in their decision making and valuation of water service changes.

Despite a desire for greater standardisation in PR14, and clarity of information presented to respondents, there was again a wide variety of approaches in detail. And the WTP surveys for PR14 showed a wide range of results. United Utilities Water (2016), referencing a report by Accent and PJM Economics (2014), quoted high, low, and median values for three water services from PR14 SP WTP surveys as examples (Table 1).

As United Utilities Water (2016) argued, whilst some differences between customer WTP values for each water service could be expected between water company areas, due to differences in factors such as household incomes and bill levels, these differences are much larger than could be explained by these factors. The high values, and possibly some of the median values, could be considered implausibly high. For example, it might be judged that it would probably be expected that most people to be willing to accept one short-term water supply interruption for much less than £1670, and probably less than £206 per year for 1 short-term interruption per year. Some of these differences in values might have been tied up with the presentation of risk to respondents and to respondents’ evaluation of these different presentations of risk. This is could be described as an issue of ‘insensitivity to scope’. It should also be noted that the coefficients from DCEs are themselves random variables. Consequently, even in the same water company customers at one particular PR, a different random sample of respondents will give rise to a different WTP value, and this different WTP value could be quite large depending on the variance and probability distribution.

The Accent and PJM Economics (2014) study covered WTP values from DCE in PR09 and PR14, across 15 water companies, for WTP to reduce a variety of water services problems: discoloration, taste and smell, interruptions to supply, drought restrictions, low water pressure, leakage, sewer flooding, odour from sewage treatment works, and pollution incidents. The Report noted the wide dispersion in WTP values for one incident (e.g. 1 property affected by 1 interruption to supply for 3–6 hours), across company studies. However, the Report did not attempt to explain differences in these WTP values.

In PR14, in addition to stated preference (SP) research on consumer preferences for water service changes, there was a broadening of the outcome measures and greater use of non-SP methods to value other water company related service attributes such as the value of reducing lead content of

| Water service                                      | High value | Low value | Median value |
|----------------------------------------------------|------------|-----------|--------------|
| 1 property affected by 1 drinking water taste and smell incident | 28,537     | 196       | 2008         |
| 1 property affected by an unexpected water supply interruption (3–6 hours) | 1670       | 50        | 206          |
| 1 property affected by 1 internal sewer flooding incident | 367,291    | 22,530    | 110,800      |

Source: United Utilities Water (2016).
water; injuries to company employees; RP measures of value for beaches and bathing water quality; and also the exploration of supply loss, or boil notices, through mitigation measures such as buying water.

The evaluation of PR14 results included a report by ICF Consulting and eftec (2017) which investigated a number of issues in preparation for PR19. These included examining whether there were any limitations in SP in collecting evidence of customers’ priorities for service deliveries, and how valid the results were in business planning. The report also investigated the extent to which SP and RP could be used to make trade-offs between short-term and long-term prices and service improvement; how SP and RP can complement each other in providing inputs into PR CBAs; and the use of RP research in the context of a monopoly industry. Also examined was the issue of presenting inflationary bill increases to customers in WTP research; and recommendations on the presentation of show card materials in SP surveys.

Ofwat encouraged water companies to think about revealed preference (RP) and other valuation methods in PR14 along with SP approaches. Whilst SP could still be used to value water quality attributes, epidemiological and value transfer was also applied to value such attributes as the reduction in the lead content of water (see Willis 2012). RP methods were used in Ireland to value the impact on recreation in Irish waterways from water quality (Breen, Curtis, and Hynes 2018). The greater use of RP and other non-SP methods to value water services continued in PR19, as greater emphasis was placed on triangulation as a method to establish the robustness of values for water services amongst each water company’s customers.

**PR19 consumer research**

PR19 saw a lot more attention given to how participants could cope with information presented to them, and in what they were being asked to do. But again water companies, and the economic consultants they employed, developed different research programme methodologies. Not unsurprisingly, there was again evident differences in the values derived for many ‘common’ attributes or water service measures, prompting the question as to whether these were legitimate differences or whether they were methodologically driven.

Insensitivity to scope issues, the desire to avoid choice complexity, and the need to present choice alternatives as clearly and simply as possible, also led in PR19 to the application of maximum difference scaling (MaxDiff). MaxDiff, and best-worst scaling (BWS), were implemented in a number of PR19 studies for different water companies (see for example Metcalfe 2021). The use of Max Diff and BWS omitted asking participants to evaluate difficult risk differences (probability and impact) and concentrated instead just on the relative importance customers attached to the attributes as they impacted directly on each respondent, e.g. by asking respondents which would have the most impact, and which would have the least impact, across a set of services such as discoloured water for a week, short-term interruption to supply, sewer flooding in a nearby public area, and persistent low water pressure. Key issues for DCE and BWS are the composition of the choice sets, or the sub-sets of items to be presented to respondents; the experimental design of the choice sets; and the type of choice sets to present to respondents; and the aggregation of individual utilities or preferences (Louviere, Flynn, and Marley 2015).

This MaxDiff, and BWS approach, was coupled in PR19 with a number of other practical developments which included yet more cognitive testing of the surveys to ensure that they were as well understood as possible; the embedding within the on-line questionnaire of devices like hover buttons to give participant support without cluttering the screen; more visual prompts, again to help with comprehension; and an array of different SP and RP tools to help provide triangulation data points.

PR19 saw a greater use of non-SP methods to estimate customer values for some specific water and waste water services. Greater consideration was given to revealed preference (RP) methodology, and avertive and mitigation approaches, by some water companies. Revealed preference (RP)
Research from PR04 to PR19 has been used to investigate methodological issues in choice modeling and in behavioural economics. These have encompassed status quo effects; asymmetry and non-linearity in benefits for gains and losses; and consistency in choice behaviour; amongst others.

Scarpa, Ferrini, and Willis (2005) used PR04 data to explore the effect of explicitly accounting for systematic differences in preferences for non-status quo alternatives in economic models. They found that alternatives offering changes from the SQ did not share the same preference structure as SQ alternatives, as found by other in the marketing, environmental, and food preference literature. Monte Carlo simulation indicated that the expected bias in estimates from ignoring the SQ is substantial, but error component specifications with SQ ACS (alternative specific constant) are efficient even when biased.

DCEs often reveal a high frequency of status quo (SQ) choices. This may signal an unwillingness of respondents to evaluate the proposed trade-offs in service levels, questioning the welfare theoretic interpretation of observed choices, and the validity of the approach for regulatory purposes. Using the methodology for DCE in the regulation of water and sewerage services in England and Wales, Lanz and Provins (2015) contributed to the understanding of SQ choices in several dimensions. They controlled for the perception of the SQ and the importance of attributes in day-to-day activities; and used a split sample design to vary both the description of the SQ and the survey administration mode (online vs. in-person). They also allowed the service attributes to both improve or deteriorate, so that the SQ is not necessarily the least-cost option; and also examined SQ choices in individual choice tasks and across all tasks to identify the determinants of serial SQ choices. Their results suggested that individual SQ choices mostly reflect preferences, and thus represent important information for the regulator. However, serial SQ choices are mainly driven by cognitive and/or contextual factors, and these responses should be analysed as part of standard validity tests.
Data from PR09 was used to investigate discrepancies between willingness to pay (WTP) and willingness to accept (WTA) in the context of a stated choice experiment. Using data on customer preferences for water services where respondents were able to both ‘sell’ and ‘buy’ the choice experiment attributes, Lanz et al. (2010) found evidence of non-linearity in the underlying utility function even though the range of attribute levels was relatively small. Their results revealed the presence of significant loss aversion in all the attributes, including price. They found the WTP–WTA schedule to be asymmetric around the current provision level and that the WTP–WTA ratio varied according to the particular provision change under consideration. Such reference point findings are of direct importance for practitioners and decision-makers using choice experiments for economic appraisal such as cost–benefit analysis, where failure to account for non-linearity in welfare estimates may significantly over- or under-state individual’s preferences for gains and avoiding losses respectively.

Research into the WTP–WTA asymmetry, and the independent valuation and summation (IVS) issue, has been built into a number of PR research studies (see for example Bormann et al. 2013). It is now generally recognised that studies with multiple DCEs to cater for the large number of water service attributes, require a ‘package’ CE, or alternatively a contingent valuation (CV) question, to derive an accurate value for the ‘package’ of measures to be implemented. Values cannot be derived from the simple addition of individual water service measures from each of the DCEs. To do so would be to overestimate the value of the investment programme to customers. But how to best implement a ‘package’ exercise remains an important issue which still requires further research.

The results of mixed logit models have been used to model customer preferences at the individual level. Scarpa, Willis, and Acutt (2004) investigated alternatives ways of modelling heterogeneity of tastes for attributes for PR04, using mixed logit (MXL), error component (EC), and latent class models (LCM). LCM do not require any assumption about the mixing variable, unlike a MXL model which requires an a priori parameter distribution (normal, log-normal, uniform, etc.). LCMs for Yorkshire Water customers, in PR04, revealed considerable heterogeneity in tastes for water service improvements across different customer segments, with large percentage of customers exhibiting negative WTP for some attributes such as river water quality. LCMs can offer insights into the heterogeneity of customer preferences that are not readily identifiable through a traditional MXL model. LCM are not seen as being particularly relevant by water companies, who require a price that can be applied to all domestic customers and not just a segment of customers.

The exception has been research linked to customers in ‘water poverty’ i.e. households who spend >3% and >5% of their income on water bills. In 2009/2010, 23.6% of households were spending more than 3% of their income on water and sewerage; and 11.5% per spending >5% of their income on water and sewerage disposal (Bradshaw and Huby 2013). With water bills increasing faster than income over the period 1999–2019, this was a formula for increasing water poverty. Researchers have either run separate DCE for vulnerable customers; or if sample size was restricted for such customers, included a dummy variable for different classes of vulnerable customer (e.g. age >55 years; ethnicity, low income, and water poverty customers). Willis, Scarpa, and Acutt (2005) showed low income customers were willing to pay less than the average income customer; and non-white customers less than white customers, for some water service improvements.

A further indication of the value of estimating individual-level preferences, individual-level data was used by Reid, Chalak, and Hecht (2010) to perform a market simulation of different possible investment programmes. Market simulation results showed that the majority of customers preferred a ‘low improvement’ scenario with a moderate associated increase in their annual bill. Virtually none of the respondents favoured a scenario where services were degraded in return for a small decrease in annual bills. This research showed how an investment optimisation tool, using CE customer surveys, can be applied in the business setting and used to identify an optimal investment plan for water companies.

Unobserved heterogeneity of error scale in choice models is a recent extension of the issue of heterogeneity of taste intensity. Thiene, Scarpa, and Louviere (2015) investigated this issue by
using choice panel data with specifications that simultaneously handle inter-personal variation in scale and taste. The aim was to separate differences in preference intensities across respondents from differences in the degree of consistency in choice behaviour, or ‘preference discrimination’, while accounting for correlation between the two. This research also investigated the issue of attribute non-attendance in choice analysis. Thiene, Scarpa, and Louviere (2015) developed a finite mixing model to simultaneously address all these issues, using stated preference data on tap water attributes in a water industry regulation study in Vicenza, Italy. However, these issues have not received any attention in UK PR studies of water service attributes.

The value of tap water quality has also been assessed through market based approaches. Lanz and Provins (2016) designed a survey to provide quantitative evidence about household demand for tap water supply attributes: water hardness and aesthetic quality in terms of taste, smell and appearance. They obtained water company customers’ expenditures on products that improve the overall experience of these attributes. For water hardness, results showed around 14% of households employed at least one water softener device or purchased products such as softening tablets or descaling agents. For the aesthetic quality of tap water, around 39% of households reported some averting behaviour, such as the use of filtering devices, purchase of bottled water, or addition of squash or cordial. An econometric model of household expenditure on these products, with level of service quality on a disaggregated regional basis measuring water hardness and aesthetic quality, suggested that households’ decisions to incur averting expenditures varies with service quality in a statistically and economically significant manner, providing novel evidence that households actively respond to non-health related aspects of tap water quality.

In further research Lanz and Provins (2017) investigated how households choose to incur averting expenditures as a substitute to the ‘public good’ provision level of tap water characteristics in terms of hardness and aesthetic quality (taste and odour). When unobserved heterogeneity affects both the perceived quality of the water and averting behaviour, identification of the demand function is affected by the problem of endogeneity. Averting expenditures and quality perception were modelled by Lanz and Provins (2017) as simultaneous processes, whereby changes in an objective measure of provision translates into an improvement in perceived quality, and improved perceived quality reduces averting expenditures.

Water companies have started using other innovative valuation tools such as reverse auctions to derive better estimates of the cost of environmental services. Reverse auctions have been used to assess the costs of reducing water pollutants, increasing biodiversity, natural flood management, and other natural capital benefits; through bids for buffer strips, arable reversion, and other environmental measures (see Claydon 2019; Balmford 2020).

The recent use of natural capital (NC) and natural capital accounting methodology in UK government environment plans (see Curnow, 2019) has shown the potential applications of the NC approach in environmental project and policy appraisal, and regulation, including water and asset management planning by water companies. The applicability of this methodology to valuing ecosystem services has been reviewed by Tinch et al. (2019); whilst Davis et al. (2019) demonstrated the use of this methodology to appraise the value of improvements to water quality needed to establish new salt marshes as a means of realigning coastal sea defences. Some water companies are adopting an ecosystem resilience and natural capital approach to managing their water catchment areas, as a way of improving water supply and water quality delivery (see United Utilities Water 2018), including reducing discoloration, reducing water pollutants, and improving waste water discharges into the environment.

Conclusions

Methods and practices, to value water and waste-water services to customers, have evolved continually since the initial price reviews following privatisation of water companies in 1989. However, there is still diversity in the detail of the approaches used by different water companies; and a wide
range in estimated customer values, for different water and waste-water services, between water companies. This heterogeneity of approaches has advantages and disadvantages. Advantages include testing different approaches, to see which works best; whilst diversity of approach reduces the risk of producing erroneous customer values across the entire industry; and allows for SP methodology to evolve in future price review appraisals. Disadvantages include a diversity, and often a wide dispersion, of customer values for specific water service measures across different water companies; and scepticism about the results.

The wide range of customer values for each water service, between different water companies, generates concern for some people in the regulatory process. Despite the tremendous thought and attention to methodology, that was incorporated in the PR19 studies, members of some water company Consumer Panels were sceptical about the process and values derived (see for example The South Staffordshire and Cambridge Water Customer Panel 2019).

Despite some reservations, SP methods will continue to form a large element in estimating customers’ values for water and waste water in the future, simply because, for some water and waste-water services, there are no other feasible methods to estimate the customers’ values for changes in the provision of these services. The ‘public good’ nature of some water services renders the use of non-SP techniques infeasible: there are simply no markets, nor weakly complementary private market goods, for some water services.

There is potential here for a cross-industry, or cross water company, DCE study, to estimate customers’ values for core services. This might allay some concerns about the diversity of values for the same service improvement between different water companies; or at least provide a comparator against which to judge local customer valuations for specific water services.

Where non-SP techniques can be applied to some water services, for example linking river and bathing water quality to recreation use or house prices, through RP methods, then future price review research is are likely to employ these more extensively. After all, they can, for some water services, be used as a triangulation method to check and validate SP values. Travel cost models (TCMs) can assess environmental attributes such as water quality, by estimating how much recreationalists are willing to pay to gain access to better water quality; whilst hedonic price models (HPMs), can be used to determine values for river water quality, bathing water quality, and odour from sewerage treatment works, by linking house prices with these variables. But RP methods are not without their own problems: multicolinearity amongst variables, omitted variable bias, and functional form in hedonic price models (HPM); and assumptions about travel and time cost, car cost sharing, multipurpose trips, omitted variables affecting site choice, substitute sites, functional form, sampling issues (visitors v trips), etc., all of which can affect travel cost method (TCM) results.

The use of natural capital and natural capital accounting also adds an additional perspective, although natural capital values are often derived using SP techniques.

SP, RP, and market based approaches, will continue to play a role in future price review research by water companies, to ensure that customers’ preferences, and values for water services, are reflected in water company business plans, as part of the regulatory framework for the water industry in the UK.

The increased diversification of valuation methods, and inclusion of qualitative methods, would give greater impetus to the increased use of triangulation in future PRs. This potentially could also enhance the understanding of why people make the choices that they do; and why valuations differ between water companies. It would also aid an assessment, and use, of ‘value transfer’ as a means of deriving accurate and reliable valuations of water services across different sets of customers facing different combinations of water supply, waste water disposal, and environmental services, in their areas.

Undoubtedly the application of stated preference (SP) in future price reviews will build upon the experience of using DCEs and CV in previous price reviews; and on the continued development and refinement of these techniques by academic researchers. The application of SP (both DCE and CV) experiments in future price review research should aim to address trust and realism issues. The
effect of trust on WTP values was raised by AECOM (2017) in a PR19 study for Yorkshire Water, although largely in terms of a literature review. The impact of trust on WTP can be quantified: Powe, Willis, and Garrod (2006), for street lighting improvements, found that a lack of trust had a significant impact of WTP values.

Asymmetry in WTP/WTA values for service changes is also likely to feature in future PRs. The effect of greater familiarity in making non-market choices has been explored by Chilton et al. (2012) and Radmehr, Willis, and Metcalf (2018). This research has shown the hypothesis that WTP equals WTA cannot be rejected for an incentivised mechanism, and it appears to control for the individual’s strategic behaviour bias as a treatment against over-estimating WTA. The continued development of DCEs and CV, should lead to greater confidence in SP results in future water price review research.

Future PRs will continue to assess customers’ preferences and values for perennial water services such as unexpected interruptions to supply, low water pressure, internal and external sewage flooding, odour from sewage treatment works, etc. But other challenges in environmental and resource economics are also likely to feature more prominently in future PRs. Bretschger and Pittel (2020) have outlined twenty key challenges for environmental economics, including carbon and climate neutrality; risk, uncertainty and resilience to environmental shocks (e.g. floods and droughts); behavioural environmental economics; equitable use of the environment (equity and fairness); loss of biodiversity and natural capital; valuing and paying for ecosystem services. Some of these have already featured in PR studies. But other have received scant attention, including equity issues which deserves substantially more consideration. Modelling the market share of customers’ has shown that a substantial number of customers are willing to pay less than the estimated mean WTP value for the set of water service improvements. In one study, 40% of customers were willing to pay less than the estimated mean WTP value for the set of environmental improvements (reducing pollution incidents, and improving river water quality, and the number of beaches with excellent bathing water quality); and 51% of customers were willing to pay less than the estimated mean WTP value for the set of waste water improvements (reducing the incidents of external sewage flooding, and internal sewage flooding, and the number of properties affected by odour from sewage treatment works). Implementing such sets of improvements could result a reduction in welfare for a significant proportion of customers. Such water service improvements are far from attaining a Lindahl equilibrium in terms of price and optimum level of provision of water supply, waste water disposal, and environmental services, for water company customers.

Note

1. Separability implies that the marginal rate of substitution (MRS) between any two goods in one consumption group is unaffected by a change in the quantity of a good in another group. Separability allows values for two or more goods in different groups to be added linearly.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Accent and PJM Economics. 2014. Comparative Review of Willingness-To-Pay Studies. London: Accent Market Research.

Aecom. 2017. PR19 Understanding Customer Values Work Package 6 – Trust Experiment. Report to Yorkshire Water. London: AECOM.

Anciaes, P. 2022. “Revealed Preference Valuation of Beach and River Water Quality in Wales.” Journal of Environmental Economics and Policy 10, 11 (1), 75–94. https://doi.org/10.1080/21606544.2020.1864778.

Anciaes, P., P. Metcalfe, and A. Sen. 2020. “A Combined SP-RP Model to Estimate the Value of Improvements in Freshwater Angling in England.” Journal of Environmental Economics and Policy 9 (2): 167–187.
Balmford, B. 2020. “Designing Auctions for the Procurement of Environmental Services: Insights from Practice.” Paper presented to the UKNEE Annual Conference, London.

Bormann, A., A. Provins, A. Chalak, L. Gahan, M. Baker, and S. Reid. 2013. “South Staffs Water PR14 Stated Preference Study: Final Report to South Staffs Water.”

Bradshaw, J., and M. Huby. 2013. “Water Poverty in England and Wales.” *Journal of Poverty and Social Justice* 21 (2): 137–148.

Breen, B., J. Curtis, and S. Hynes. 2018. “Water Quality and Recreational use of Public Water Ways.” *Journal of Environmental Economics and Policy* 7 (1): 1–15.

Bretscher, L., and K. Pittel. 2020. “Twenty Key Challenges in Environmental and Resource Economics.” *Environmental and Resource Economics* 77: 725–750.

Burge, P., F. Tsang, and C. Rohr. 2007. *Northumbrian Water Customer Priorities Research: Modelling Report*. Cambridge: Rand Europe.

Cascade Consulting and eftec. 2011. *The Use of Revealed Customer Behaviour in Future Price Limits*. Report to Ofwat. Manchester: Cascade Consulting, London: eftec.

Cascade Consulting, eftec, and ICS Consulting. 2010. *Review of Cost-Benefit Analysis and Benefit Valuation*. Report to UK Water Industry Research Ltd., London. Report Ref. No. 10/rg/07/18. Appendix 3 Quantitative Analysis. Manchester: Cascade Consulting.

Chilton, S., M. J. Lee, R. McDonald, and H. Metcalf. 2012. “Does the WTA/WTP Ratio Diminish as the Severity of a Health Complaint Is Reduced? Testing for Smoothness of the Underlying Utility of Wealth Function.” *Journal of Risk and Uncertainty* 45 (1): 1–24.

Claydon, E. 2019. “Reverse Auctions as a Tool for Securing Multiple Environmental Outcomes.” Paper presented to the UKNEE Annual Conference, London.

Curnow, J. 2019. “Environmental Economics in UK Environmental Policy: Defra’s 25 Year Environment Plan.” *Journal of Environmental Economics and Policy* 8 (4): 353–358.

Davis, K. J., A. Binner, A. Bell, B. Day, T. Poete, S. Rees, G. Smith, K. Wilson, and I. Bateman. 2019. “A Generalisable Integrated Natural Capital Methodology for Targeting Investment in Coastal Defence.” *Journal of Environmental Economics and Policy* 8 (4): 429–446.

DeShazo, J. R., and G. Fermo. 2002. “Designing Choice Sets for Started Preference Methods: The Effects of Complexity on Choice Consistency.” *Journal of Environmental Economics and Management* 44: 123–143.

Eftec. 2009. *Development of Guidelines for Use of Benefits Transfer in Policy and Project Appraisal*. Report to Defra. London: Eftec.

Environment Agency. 2003a. *Economic Appraisal for the Environment Programme in PR04*. Bristol: EA.

Environment Agency. 2003b. *AMP4 Benefits Assessment Guidance—Corrigenda*. Bristol: EA.

Foundation for Water Research. 1994. *Assessing the Benefits of River Water Quality Improvements*. Marlow: FWR.

Hanley, N., R. E. Wright, and B. Alvarez-Farizo. 2006. “Estimating the Economic Value of Improvements in River Ecology Using Choice Experiments: An Application to the Water Framework Directive.” *Journal of Environmental Management* 78 (2): 183.

Hensher, D., N. Shore, and K. Train. 2005. “Households’ Willingness-to-Pay for Water Service Attributes.” *Environmental and Resource Economics* 32: 509–531.

HM Treasury. 2020. *The Green Book: Central Government Guidance on Appraisal and Valuation*. London: HM Treasury.

Hoehn, J. P., and J. B. Loomis. 1993. “Substitution Effects in the Valuation of Multiple Environmental Programs.” *Journal of Environmental Economics and Management* 25 (1): 56–75.

Hoehn, J. P., and A. Randall. 1989. “To Many Proposals Pass the Benefit Cost Test.” *American Economic Review* 79: 544–551.

ICF Consulting and eftec. 2017. *Improving Willingness-to-Pay Research in the Water Sector*. Report to the Consumer Council for Water, Birmingham. London: ICF Consulting.

Lanz, B., and A. Provins. 2015. “Using Discrete Choice Experiments to Regulate the Provision of Water Services: Do Status Quo Choices Reflect Preferences?” *Journal of Regulatory Economics* 47 (3): 300–324.

Lanz, B., and A. Provins. 2016. “The Demand for Tap Water Quality: Survey Evidence on Water Hardness and Aesthetic Quality.” *Water Resources and Economics* 16: 52–63.

Lanz, B., and A. Provins. 2017. “Using Averting Expenditures to Estimate the Demand for Public Goods: Combining Objective and Perceived Quality.” *Resource and Energy Economics* 47: 20–35.

Lanz, B., A. Provins, I. Bateman, R. Scarpa, K. Willis, and E. Ozdemiroglu. 2010. “Investigating Willingness to Pay – Willingness to Accept Asymmetry in Choice Experiments.” In *Choice Modelling: the State-of-the-Art and the State-of-Practice: proceedings from the Inaugural International Choice Modelling Conference*, edited by S. Hess and A. Daly, 517–542. Bingley: Emerald.

Louviere, J. J., T. N. Flynn, and A. A. J. Marley. 2015. *Best-Worst Scaling: Theory, Methods and Applications*. Cambridge: Cambridge University Press.

Metcalfe, P. J. 2021. “Max Diff and Contingent Valuation: Welsh Water PR19.” *Journal of Environmental Economics and Policy* 10.
NERA. 2011. Carrying Out Willingness to Pay Surveys. Report to Ofwat. London: National Economic Research Associates (NERA).

OFWAT. 2000. Maintaining Serviceability to Customers. MD161: Letters to Managing Directors of All Water and Sewerage Companies and Water Only Companies. Birmingham: Office of Water Services.

Palmer, Ronan. 1998. Multi Attribute Techniques and Their Application to Water. Bristol: Environment Agency.

Powe, N. A., K. G. Willis, and G. D. Garrod. 2006. “Difficulties in Valuing Street Light Improvements: Trust, Surprise, and Bound Effects.” Applied Economics 38: 371–381.

Radmehr, M., K. Willis, and H. Metcalf. 2018. “A Mechanism to Derive More Truthful Willingness to Accept Values for Renewable Energy Systems.” Heliyon. doi:10.1016/j.heliyon.2018.e00503.

Reid, S., A. Chalak, and J. Hecht. 2010. “Determining the Optimal Investment Plan for Water Utilities: the Case of Veolia Water Central.” Water Supply 10 (3): 367–375.

Scarpa, R., S. Ferrini, and K. Willis. 2005. “Performance of Error Component Models for Status Quo Effects in Choice Experiments.” In Applications of Simulation Methods in Environmental and Resource Economics, edited by R. Scarpa and A. Alberini, 247–273. Dordrecht: Springer.

Scarpa, R., K. G. Willis, and M. Acutt. 2004. “Individual-Specific Welfare Measures for Public Goods: A Latent Class Approach to Residential Customers of Yorkshire Water.” In Econometrics Informing Natural Resources Management, edited by P. Kouanddouri, 316–337. Cheltenham: Edward Elgar.

Scarpa, R., K. G. Willis, and M. Acutt. 2007. “Valuing Externalities from Water Supply: Status Quo, Choice Complexity and Individual Random Effects in Panel Kernel Logit Analysis of Choice Experiments.” Journal of Environmental Planning and Management 50 (4): 449–466.

South Staffordshire and Cambridge Water Customer Panel. 2019. “Independent Report to OFWAT to Accompany the PR19 Business Plan of South Staffs Water (SSC).”

Swait, J., and W. Adamowicz. 1999. Choice Environment, Market Complexity, and Consumer Behavior: A Theoretical and Empirical Approach for Incorporating Decision Complexity into Models of Consumer Choice. Staff Paper 99-04. Edmonton: Department for Rural Economy, University of Alberta.

Thiene, M., R. Scarpa, and J. J. Louviere. 2015. “Addressing Preference Heterogeneity, Multiple Scales and Attribute Attendance with a Correlated Finite Mixing Model of Tap Water Choice.” Environmental and Resource Economics 62: 637–656.

Tinch, R., N. Beaumont, T. Sunderland, E. Ozdemiroglu, D. Barton, C. Bowe, T. Börger, et al. 2019. “Economic Valuation of Ecosystem Goods and Services: A Review for Decision Makers.” Journal of Environmental Economics and Policy 8 (4): 359–378.

United Utilities Water. 2016. Improving Customer Research and Engagement. Warrington: United Utilities Water.

United Utilities Water Ltd. 2018. Ecosystem Resilience: Ecosystem Resilience Through Catchment Management, Partnership and Markets. Chapter 4: Supplementary Document. Document Reference S4004. Warrington: United Utilities Water.

Willis, K. G. 2008. “Evaluating Environmental Benefits from Changes in Water Abstraction and Waste Water Disposal.” Water Policy 10 (5): 439–458.

Willis, K. G. 2012. Valuing Health Benefits from Reductions in the Lead Content of Tap Water: A Value Transfer Analysis. Report to Yorkshire Water. Newcastle upon Tyne: Centre for Research in Environmental Appraisal and Management, University of Newcastle.

Willis, K. G., R. Scarpa, and M. Acutt. 2005. “Assessing Water Company Customer Preferences and Willingness-to-Pay for Service Improvements: A Stated Choice Analysis.” Water Resources Research 41: W02019.