

## Comparing the Results of Economic Valuation Studies - *the case of using stated preference techniques for valuing improvements to the water environment*

The number of studies applying stated preference techniques in the context of policy and projects about the water environment is increasing. Seemingly similar studies are producing widely different results. An informed interpretation of these results requires an understanding of the key parameters that influence them.

This briefing aims to summarise these key parameters and their implications using the recent studies on the economic benefit of implementing the Water Framework Directive and other environmental improvements planned by water companies in England as examples.

### Introduction

Economic valuation refers to the use of monetary units to measure individuals' preferences for goods and services. Several valuation techniques, which use different types of economic data, exist. But, in essence, all try to estimate the change in human wellbeing (expressed through preferences) arising from the changes in the provision of goods and services.

Individuals demonstrate their preferences for environmental goods and services by their behaviour as customers in actual markets (e.g. purchasing bottled water); by their behaviour in so called surrogate markets (e.g. spending for the recreational activities associated with the water environment) and by stating their

preferences directly through surveys. This briefing note concerns the latter: *stated preference techniques*.

Stated preference techniques use questionnaires to create hypothetical markets within which individuals can trade off changes in the provision of goods and services against their income for or against such changes. The questionnaires elicit respondents' *willingness to pay* (WTP) money to secure a positive change or to prevent a negative one; or *willingness to accept* (money as) *compensation* (WTA) to forgo a positive change or to tolerate a negative one.

There could be several motivations for individuals stating positive WTP or WTA related to: (i) use values - people making

use of the good or service on offer, or indirectly benefiting from them as in the case of some ecosystem services - and/or; (ii) non-use values - people caring for others using the good or service now (altruism) and in the future (bequest) and/or caring for the goods and services for their own sake (existence).

A variety of introductory and advanced texts are available regarding valuation techniques (see Further Reading). This briefing assumes a general but not advanced knowledge on the subject and focuses on the key parameters that help decision-makers interpret the results of stated preference studies and understand the origin and meaning of differences between study results. The likely effects of key parameters are illustrated by comparing the Water Framework Directive (WFD) benefits study conducted for Defra (2007) and valuation studies conducted for water companies in England as part of the 2009 Price Review (PR09) (see boxes for comparisons and Further Reading for references). Readers are assumed to be familiar with these studies.

## Key Parameters for Comparison

Key parameters for comparison include:

- The purpose of the study;
- The valuation methodology;
- The affected population;
- The good (and/or service);
- The baseline;
- The change;
- The payment vehicle, and
- The validity of the results.

## The purpose

Economic value evidence can be used for a variety of decision making purposes including:

- Cost benefit analysis for policies, programmes and projects;
- Assessing the importance of an issue;
- Setting priorities within a sector or across sectors;
- Establishing the basis for an environmental tax or charge;
- Green national income or corporate accounting;
- Assessment of liability for environmental damage, and
- Estimating empirical discount rates.

These and other purposes affect the way in which the stated preference questionnaire is designed, in turn influencing the parameters discussed here and hence the results.

The purpose of the [WFD benefits](#) study is to estimate the economic value of the benefits of national and regional compliance with the 'good ecological status' criterion defined under the WFD. Such benefit estimates can be used within cost benefit analyses at both the national level and River Basin District level. The study also looks at the timing of WFD compliance (compliance vs. time derogation).

Typically [water company](#) PR09 studies estimate the benefit of environmental improvements amongst other services provided to customers. The main purpose is to identify priorities between different services and contribute to cost benefit

analysis and investment strategies at the company level.

The difference in the purpose reflects the differences in the decision-making context, the good, baseline and change that are valued and determines the choice of methodology and other key parameters. And hence difference in purpose influences relative results.

### *The methodology*

There are two main methodologies within the stated preference techniques. **Contingent valuation** (CV) presents the good or service as a bundle of its characteristics (or attributes) and requires the respondents to ‘take’ or ‘leave’ the entire bundle. The trade off is between the provision of the bundle and the price of its provision. The WTP or WTA question can be asked in different so-called ‘elicitation formats’. The two that are relevant here are: (i) dichotomous choice (e.g. “are you willing to pay £x?” with the amount £x being varied across a sample of respondents to allow the analyst to estimate average values) which may be asked once or repeated several times depending on the responses, and (ii) payment card which shows different money amounts for the bundle for which respondents are asked to state acceptance, rejection or uncertainty. The results are expressed in terms of £ per household per year for the entire bundle described.

**Choice modelling** (CM; also referred to as choice experiments; CE), on the other hand, presents the good or service not as a bundle but in terms of its constituent attributes (e.g. water based recreation could be described in terms of clarity of

water, fish catch rate, entry fee to site, etc.). Each of these attributes can be set to different quality or quantity levels. Setting levels for each attribute defines a given state. Survey respondents then choose between two or more such states (an operation which is then repeated with different combinations of attribute levels to generate more information on respondent preferences). By including price in the exercise analysts can observe the trade off between money and different levels of a given attribute in terms of £ per household per year per unit change in that attribute.

Note that the ‘price’ (often referred to as the ‘bid level’) in both CV and CM exercises does not necessarily equate to the actual cost of provision. Several different estimates of the price are tested in order to have sufficient variation in the responses to enable statistical analysis. Thus, the results of CV and CM studies do not measure approval of the cost of providing goods and services, but measure the benefits of such provision.

Because of the different formulations of goods presented in CV and CM exercises care needs to be taken to conduct any controlled comparison of results. In the few cases where this has been done the findings from such comparisons are mixed. In some cases values have been similar, while others have shown a tendency for CM to yield higher values than CV. Even within CV studies, different elicitation formats could yield different results (e.g. payment card estimates are expected to be smaller than those from a dichotomous question). Differences here are generally not a problem but simply an outcome of different approaches. The reasons for the

differences are several and the discussion is too involved to cover here in full. Further reading at the end of the briefing is a good start for a more detailed discussion.

The way the chosen methodology is administered ('survey mode') may also affect results. Face-to-face interviews with respondents are typically preferred (especially for complex issues or questionnaires involving visual materials) to telephone, mail or online surveys. The latter two modes, in particular, are more likely to have biased samples (through self selection of respondents with particular interests in the topic) and low response rates.

The [WFD benefits](#) study uses both CV and CE administered through face-to-face interviews. In the CV version, both dichotomous choice and payment card are employed. The results show that (all other things being equal), the CE implies larger WTP than the CV, and that the dichotomous choice format yields larger estimates than the payment card.

Some [water company](#) PR09 studies use both CV and CM while others use only CM, again through face-to-face interviews. For these studies, the WTP derived from the CM are usually larger than those derived in the CV (dichotomous choice) question.

Without a more involved comparison, it is not possible to identify any effect the choice of method may have on the relative results.

### *The affected population*

Stated preference surveys should be conducted with representative samples of the population affected by the changes of concern. Since different groups of individuals (e.g. users and non-users) are likely to have different preferences, the definition of the population affected and representativeness of the sample are crucial parameters influencing the valuation results.

The affected population for the [WFD benefits](#) study is the national population.

The affected population for the [water company](#) PR09 studies is the company customer (household and business) base.

All studies have large samples and are satisfactorily representative of the relevant affected population.

In general, the larger the population affected (implying larger goods), the more likely that they will have greater altruistic and other non-use values that could add to the WTP for a larger good. Larger populations also lead of course to larger aggregate benefits. Thus, the larger population in the WFD benefits study is clearly linked to the larger result of the study.

### *The good*

The type and size of the good or service valued (together with the definition of baseline and the change) determine the size of the value estimate. Broad definitions like 'water quality' or 'water supply' are likely to be too general to explain the differences between the results of studies. Information on the attributes of

the good, whether expressed separately as in CM or as part of a bundle as in CV is also needed.

The good in the [WFD benefits](#) study is *all* water bodies in a 30 mile radius (local) and nationally (allowing for separate identification). Water bodies include lakes, rivers, estuaries and the coast as defined by the WFD. The questionnaire presents detailed definitions of what quantity, morphology and quality changes are involved. Attributes separately described in the CM version are the level of low, medium and high quality at local and national scale now, in 2015 and in 2027.

The good described in the [water company](#) PR09 studies are the water and wastewater services provided by the companies. Within this, attributes that relate to aspects of the water environment include: percentage of river length complying with the current quality standard of the Environment Agency (RE classes); compliance with bathing water quality standards; and number of category 1 or 2 pollution incidents per year. These apply to the local (water company) area only and vary with respect to comparability with the WFD benefits study. Definition of water quality elements in each of these attributes is necessarily brief and covers the next investment period.

As far as the quality of water bodies is concerned, the WFD study describes a larger good than that encompassed by the environmental attributes of the water company studies. This is likely to be one of the reasons why the former estimates higher values.

### *The baseline*

The changes in the provision of the good are described relative to a baseline. The baseline is the option of ‘doing nothing’ or doing nothing further than what is already planned or what needs to be done to continue to maintain current quality and quantity of provision. The baseline can thus be static (as is ‘now’) or dynamic (e.g. the situation in 2010 if all already committed investments but no other action are undertaken, or other variants which account for changes in provision over time not associated with the action in question).

The baseline in the [WFD benefits](#) study is ‘as is now’ (no deterioration and no improvement) which does not include the effects of non-WFD improvements between now and 2015 (or 2027).

The baseline in the [water company](#) PR09 studies will typically include all planned investments and hence the baseline is the future state of the attributes.

Thus, there is a difference in the way baseline is defined between the two studies, with the former measuring changes from a worse baseline than the latter.

There is also an ongoing discussion about the appropriateness of different baseline data in terms of the definition of ‘good quality’. In the WFD benefits study, baselines for the local areas vary according to the local data. The baseline for all water bodies in the national part of the WFD benefits study is 15% compliance with Directive criteria.

In the water company studies, baseline percentage of compliance with the current

RE criteria is much higher (e.g. at 84% for one company). This is most likely to contribute to the higher WTP results from the WFD benefits study.

Note that the baselines tend to be different amongst water company studies which also lead to differences amongst the results.

Baseline descriptions are scientific inputs to valuation studies and hence economic analysis of the responses cannot conclude whether these descriptions were right or wrong. This discussion is simply to point out the differences.

### *The change*

The change in the provision of good and service could be qualitative, quantitative, or both. The change could affect the entire bundle of attributes in the same way (e.g. all attributes improving) or different attributes can change in different ways.

In any case, the change traded off in economic valuation is marginal (e.g. changes in water quality, not the value of the entire water environment). The relative change is important as well as the absolute change. For example, a 1% improvement in water quality that shifts a water body above the good quality threshold (however this threshold is defined) may be valued much more than a 1% improvement above that threshold (or possibly even below it).

The change in provision described within the [WFD benefits](#) study is an improvement of water bodies from the current baseline to that qualifying under WFD good ecological quality standards. These are, in most cases, more stringent than RE classes.

The quality standards are described in showcards with the change expressed in percentage of water bodies qualifying under different scenarios.

The change in the [water company](#) PR09 studies is defined in terms of compliance with existing standards (such as RE classes or bathing water quality) which are more specific and less encompassing than expectations of good ecological quality status under the WFD.

In both sets of studies, the change is an improvement with the exception of some water company studies also investigating WTP to avoid degradation from the baseline which are generally separately analysed. Because of presenting a larger good and a worse baseline, the change valued in the WFD benefits study is larger than in most water company studies. This points to the higher value result obtained in the former.

### *The payment vehicle*

Most stated preference studies define the method by which payment (or compensation) is to be made. This is known as the 'payment vehicle'. The appropriate vehicle is selected according to the institutional context of the delivery of the good or service and can be taxes, bills, general price increases, entry fees etc.

The [WFD benefits](#) study uses both water bill increases and increases in other household payments as the payment vehicle. This selection reflects the institutional setting in which the WFD will be delivered and is also linked to the respondents valuing this delivery as a

member of national population as well as water company customers.

[Water company](#) PR09 studies use water bill increases alone and this links the trade-off to one of customer preferences and company priorities.

While different payment vehicles may affect the protest rate (see below), there is little research regarding the impact of payment vehicle choice upon WTP results - the important point being that the appropriate vehicle is used, which is the case in both sets of studies.

### *Validity of results*

Validity refers to the degree to which a study measures the intended economic value. Since it is generally not possible to test the results of a stated preference study against behaviour in actual markets, some internal validity tests are typically designed into the questionnaire (e.g. asking the same question in different ways or comparing attitude responses to valuation responses). Validity of WTP (or WTA) is also tested by analysing the factors determining the estimates against prior expectations based on economic theory and empirical evidence (e.g. income should be a significant and positive determinant meaning that all else equal, higher income should correspond to higher WTP).

There are several reasons why WTP estimates may be 'biased'. Strategic behaviour may lead respondents to give valuations higher or lower than their true preferences for the change described (e.g. they may state a value to support a 'good cause' (warm glow) without considering the specific details of the scenario).

Alternatively some respondents may protest against certain aspects of the questionnaire and give zero WTP responses despite holding positive values.

The stated preference studies are designed and tested in order to avoid such biases. But, if they do occur, the way they are handled in data analysis and the overall statistical model used to analyse the responses could also have a significant effect on the results. The crucial point is that the model selected is the best fitting the data and that the biases are identified and handled according to best practice.

[WFD benefits](#) study has exceptionally low rate of protest responses at 4%. Different scenarios test the sensitivity of responses to different levels of changes in quality - such sensitivity shows the warm glow (simply giving for a good cause) is not significant. In any case, in a study like this that presents a significant improvement in a very large good (all water bodies) it would be hard to separate 'warm glow' from genuinely high WTP responses. Other prior expectations such as a significant and positive income effect are met.

[Water company](#) PR09 studies, in general, have higher protest rates which is typical of other studies in the sector. Prior expectations based on economic theory are also met.

On the basis of what is available in the study reports, there are no apparent statistical analysis related reasons that can explain the difference between the two sets of studies.

## Conclusions

This briefing discusses the key parameters that can help explaining the difference between the results of seemingly similar studies. The differences in these parameters show that the studies are in fact not similar. The box here summarises the discussions above.

In comparison to water company PR09 studies, [WFD benefits](#) study measures a bigger change (from a worse baseline) in the quality of a larger good (all water bodies) that affects a wider population (national). These factors are expected to lead to higher WTP estimates than the water company studies.

[Water company](#) PR09 studies, in turn, measure smaller changes (from a better baseline) in the quality of smaller goods (river, bathing water, pollution incidents separately) that affect smaller populations (local company customers). These factors are expected to lead to lower WTP estimates than the WFD benefits study.

## Further Reading

WFD benefits study:

NERA and Accent (2007) *The Benefits of Water Framework Directive Programmes of Measures in England and Wales*, report to Defra, London.

Water company studies:

These are not (at least not yet) publicly available. Our assessment is based on our first hand knowledge of these studies for United Utilities and Southern Water, and our reading of the study for Thames Water.

Economic valuation (introductory level):

Bateman, I.J. (2007) Valuing Preferences Regarding Environmental Change, in Jules Pretty, et al., (Eds.) *The SAGE Handbook of Environment and Society*, Sage, London, pp.155-171.

Economic valuation (more advanced):

Bateman, I.J., Carson, R.T., Day, B., Hanemann, M. Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Ozdemiroglu, E., Pearce D.W., Sugden, R and Swanson, J. (2002) *Economic Valuation with Stated Preference Techniques: A Manual*, Department for Transport (UK) and Edward Elgar: Cheltenham.

**This briefing is prepared by eftec and reflects the views of the authors only.**

**For further information about eftec, please visit**  
[www.eftec.co.uk](http://www.eftec.co.uk)