

Guidelines for the application of Stated Preference methods

Stated Preference (SP) methods are survey-based approaches to valuing non-market goods and services. A sample of the reference population is asked to complete a questionnaire where a good or service is offered in a hypothetical market, for example, the improvements in water quality or biodiversity associated with an intervention policy. Willingness to pay (WTP) for the good is then estimated using statistical procedures.

The main advantage of SP methods over Revealed Preference (RP) methods is that they allow the estimation of non-use and option values. Their main disadvantage is hypothetical bias since respondents state their preferences in a hypothetical market (Boyle, 2017). For example, respondents may provide an inaccurate WTP of the hypothetical goods/services due to a lack of understanding or familiarity.

SP methods are categorized into Contingent Valuation (CV) and Discrete Choice Experiments (DCE). In CV surveys, people are asked to state their willingness to pay (WTP) or willingness to accept (WTA) for a hypothetical change in the provision of a good (Mitchell & Carson, 1989). In DCE surveys, respondents are asked to make choices over goods or policies, defined in terms of attributes and levels in hypothetical alternative scenarios (Henser et al., 2005). That is, CV considers the good (e.g., the non-market value of a beach) as a whole and DCE as a “bundle of attributes” (Riera et al., 2012).

Designing a stated-preference survey is a lengthy and complex process. The validity and reliability of results depend on how carefully the creation process is carried out. Several basic steps should be followed, and several issues should be carefully considered (Riera et al., 2012; Johnston et al., 2017). The basic steps to follow in conducting a stated preference survey are the following (Boyle, 2017):

1. Identify the change(s) in quantity or quality to be valued, which implies identifying the current provision of the good (the status quo) and the future one. This information is included in the valuation scenario.
2. Identify whose values are to be estimated, i.e., the affected or reference population.
3. Select a data collection mode: Data can be collected by self-administered or interviewer-administered surveys. Self-administered surveys are less expensive than interviewer-administered surveys. However, they can increase self-selection bias since respondents can review the survey materials to determine if they are interested in the topic and then decide whether to participate. They require respondents to be competent readers of the language. In interviewer-administered surveys, the interviewer can provide clarification if the survey is complex; however, this may influence responses. Today's most used data collection modes are face-to-face, e-mail, or internet-based surveys. E-surveys are gaining popularity, particularly when relying on survey companies that provide a representative sample of the target population at a relatively low cost and in a short timeframe. On the other hand, face-to-face surveys allow for the administration of more complex and longer questionnaires.
4. Choose a sample size: A sample size is a matter of choosing an acceptable level of precision of the estimates within a given budget. As a general indication, a sample of 300-500 respondents can be considered acceptable; however, the sample size depends on the dimension of the reference population.
5. Design the information component of the survey instrument:

- a. Describe the item to be valued: A neutral and fair description of the item being valued, the current or baseline conditions, and the new conditions must be included. This information can be presented in written form and accompanied by visual stimuli (e.g., graphs or pictures). The information included must be tested in focus groups and pre-tests, as it affects stated preferences and influences the validity and reliability of the estimates.
 - b. Select a provisioning mechanism: It is necessary to inform respondents on how the change to be valued will be implemented (e.g., establishing protected areas to conserve the biodiversity of a beach).
 - c. Select a payment vehicle: Examples of payment vehicles are an entry ticket to a beach, parking fees, taxes, higher prices for water or electricity bills, or accommodation facilities, and voluntary donations. The choice of a payment vehicle requires balancing realism against the rejection of payment vehicles. For example, an entry ticket can be employed for marine outdoor recreation.
 - d. Select a decision rule: An example of a decision rule is that the good will be provided if aggregate benefits (aggregate WTP) exceed the cost of the project or the policy.
 - e. Select a time frame of payment: one-time payment (lump-sum), annual payments over the lifetime of the project (e.g., 20 years), payments every time the good is used (entry ticket), or perpetuity, can be considered, depending on the good to be valued.
 - f. Include substitutes and budget constraint reminders: Respondents should be reminded about substitutes and that they could spend their money otherwise.
6. Design the valuation question.
 - a. Select a response format. Response formats differ between CV and DCE. They will be discussed in the following paragraphs, which outline specific factors for each method.
 - b. Allow for values of 0: Respondents must be allowed to state they are unwilling to pay for the good.
 - c. Address protest and other misleading responses: zero value can be a real zero or a protest response. People may truly hold a zero value if they do not like the provided good and are not willing to pay for it, or they may give a response of €0 because they reject some component of the valuation scenario. A follow-up question is typically included after the valuation question to examine the reason for a zero answer. Protest responses are excluded from the data since they are not true zeros, while real zeros are retained. Strategic behavior occurs when respondents do not answer the valuation questions truthfully because they believe they can influence the survey outcome by providing different answers (Mariel et al., 2021). Questions can be included to detect strategic behaviors, or statistical routines can be used to identify responses that have an undue influence on results.
 7. Develop auxiliary questions for statistical analyses of valuation responses: Demographic questions and attitudinal questions are included in the questionnaire to collect data to be used in the analysis of responses to the valuation question (i.e., explaining the heterogeneity of preferences) and to provide information to support decision-making.
 8. Pre-test and implement the survey: Focus groups, cognitive interviews (one-on-one interviews), and pre-tests are used to test the questionnaire and avoid pitfalls such as incorrect interpretation of the information, unwanted clues to respondents, and rejection of information.

It is crucial to translate expert language into understandable and valuable information for respondents (Mariel et al., 2021). Survey pre-testing is a fundamental prerequisite for a proper survey design (Mitchell & Carson, 1989; Johnston et al., 2017).

9. Analyze data: The process of data analysis varies with the question format used.
10. Report study: The report should include all the details on how each previous step was addressed.

One further critical point to address is hypothetical bias, which refers to the difference between hypothetical and real values. This bias often leads to overestimating WTP (Bishop & Boyle, 2017). Several protocols have been proposed to remove overestimation. Ex-post, certainty scale calibration approach allows respondents to express his/her confidence or certainty about their WTP (Champ et al., 2009). Otherwise, statistical procedures are used, like the lower-bound nonparametric estimator of WTP (Haab & McConnell, 1997). Cheap talk, consequentiality, real talk, and opt-out reminder are the most ex-ante-used approaches to mitigate hypothetical bias. Cheap talk (Cummings & Taylor, 1999) is a script that aims at making respondents aware of the hypothetical bias and its consequences on estimated WTP. The consequentiality script (Carson & Groves, 2007) informs respondents that their answers can influence the policy being investigated. A follow-up question to investigate the perception of consequentiality is commonly used. Real talk (Alfnes et al., 2010; Moser et al., 2014) and opt-out reminder (Ladenburg & Olsen, 2014) have been used in DCE. Real talk informs respondents that after the hypothetical choices, they will face a real choice situation where they must purchase the good. The opt-out reminder is a script that, just before each single-choice card, reminds respondents to choose the opt-out alternative (the status quo) if they find the hypothetical alternatives too expensive.

The following paragraphs outline some key factors that distinguish each method.

Specific Factors for Contingent Valuation

Different elicitation questions are used in CV surveys. They differ in their incentive properties, the information collected from respondents, the sample size needed, and the econometric approach (Vossler & Holladay, 2018).

The open-ended question (Hammack & Brown, 1974) directly asks respondents how much they are willing to pay for the specified change without offering any cue (For instance, what is the maximum amount of money you are willing to pay for XXX?). This format is not considered incentive-compatible since it incentivizes strategic behaviors and protest responses, primarily because it is cognitively burdensome for respondents. The open-ended question can be used when respondents are familiar with the good under evaluation (e.g., fishing). In the bidding game question (Davis, 1963), respondents are asked if they are willing to pay a specific amount (starting bid). If respondents answer “yes”, the bid is increased until they say “no”; if they say “no”, it is decreased until they say “yes”. The primary disadvantage of this question format is the anchoring effect, which occurs when the respondent's initial response is influenced by the first proposed bid (starting point bias). With the payment card question (Mitchell & Carson, 1981), a list of values is provided, and respondents circle the highest amount they are willing to pay. This question format avoids starting-point bias but may still induce anchoring effects on the bids presented in the card, notably the maximum bid (Notaro & De Salvo, 2010). The dichotomous-choice question (Bishop & Heberlein, 1979) is considered the most incentive-compatible question format because it mimics real market transactions and is easy to answer. Respondents answer ‘yes’ or ‘no’ to a specific amount offered. Nevertheless, this question format suffers from the

so-called “yea-saying”: people tend to say “yes” to the offered bid amount. Since this format only provides information on whether the WTP of each individual is above or below a given amount, the double-bounded question (Hanemann et al., 1991) has been introduced. With the double-bounded question, respondents who say “yes” to the offered amount are asked a higher amount, while those who say “no” are asked a lower amount. With the multiple-bounded question format (Welsh & Poe, 1998), respondents are presented with different bid amounts, and the level of certainty they would be willing to pay is asked for each amount (definitely yes, probably yes, not sure, probably no, definitely no). For example, Ramajo-Hernandez and Saz-Salazar (2012) used a dichotomous-choice question format preceded by a 0-value screen question:

Question 1: Considering that (1) the implementation of the proposed change in water quality costs money, (2) you already pay towards some water improvements as a part of your water bill, and (3) any additional money you would pay to improve water quality will not be available to you for other purchases, would you be willing to contribute financially to such a project by means of an increase in your water bill? Yes, No, Do not know.

Question 2 (asked only to those respondents who answered “yes” to the previous question): Considering your willingness to pay, would you pay per month an extra amount of €A in your current water bill in order to enjoy the proposed improvement in water quality? Yes, No, Do not know.

A bid design is required for all the illustrated question formats, except for open-ended questions. Focus groups and pre-tests can be used to determine the bid amounts.

Specific Factors for Discrete Choice Experiments

DCE allows for estimating preferences and willingness to pay for each characteristic (attribute) of the good being valued. With CV, we can estimate, for example, the overall non-market value of a beach, whereas with DCE, we can value every single characteristic of the beach, such as the landscape, biodiversity, recreational value, and the values of each attribute that a beach possesses. At the same time, we can evaluate the beach's overall value by summing up the value of each attribute and the result of every possible beach management policy by combining the values of different attributes.

The DCE questionnaire differs from the CV questionnaire only in the way stated preferences are asked. In DCE, respondents choose between alternatives (policy scenarios) shown in a sequence of choice cards. Each alternative presents different combinations of characteristics (attributes) of the good being evaluated with different levels of achievement. In the environmental field, these alternatives typically represent the environmental changes due to intervention policies. Figure 4.5 illustrates the choice card example, taken from Börger et al. (2020). In this choice card, there are three alternatives (options A, B, and C).

	Option A (business as usual)	Option B	Option C
Seabed condition Area of seabed restored to a natural condition 	35% of area	50% of area	35% of area
Coastal communities Jobs in traditional industries lost in 10 years 	10% of jobs	25% of jobs	10% of jobs
Seabird populations in the plan area with healthy population 	66% of species	66% of species	75% of species
Additional tax For your household per year 	£0	£20	£100

Figure 1. An example of a choice card (Börger et al., 2020)

Respondents choose the preferred alternative (scenario) among those proposed, i.e., the one that maximizes their utility function. One of the scenarios typically represents the current situation, i.e., the status quo. However, it can also represent the business-as-usual situation, showing what will happen if the project or the policy is not implemented. The status quo is, therefore, generally presented with a zero-price level. In the choice card shown in Figure 1, option A is the business-as-usual situation.

A key issue is identifying attributes and levels for the choice cards. Attributes must inform respondents about environmental change and provide the decision-maker with useful information for adopting the most appropriate policy from a social perspective. It is essential that attributes and levels are easily understood by respondents, are realistic, and are relevant to them. To identify attributes and levels, interviews with experts and focus groups involving potential users of the good are conducted, and the questionnaire is pre-tested. Deciding on the number of attributes, levels, alternatives, and choice cards is also important, since as the number of these elements increases, so does the cognitive effort of respondents. The choice card with three alternatives is the most used in empirical studies, while presenting 9 - 12 choice cards is considered optimal (Caussade et al., 2005). Scarpa et al. (2011) showed that the efficiency of the response increases gradually in the first 9-11 choice cards because of a learning effect and then declines quite rapidly due to the fatigue effect.

Levels of attributes are combined in the choice cards through an experimental design. Experimental designs are mainly divided into full factorial and fractional designs. Full factorial designs are orthogonal designs - a mathematical condition that requires that the attributes must be statistically independent - and show all possible combinations of attributes and levels. Fractional designs are generated with a subset of combinations. Fractional designs can be based on the principle of orthogonality or efficiency. Efficient designs enable the obtaining of statistically significant estimates with a smaller sample size, thereby decreasing research costs (Rose & Bliemer, 2009). To reduce the number of choice cards

presented to respondents, the design can be divided into blocks, so that each respondent is only presented with a subsample of choice cards.

In Börger et al. (2020), there are four attributes, each with three levels. They use the software Ngene to generate a D-efficient design, which consists of 18 choice cards with two blocks. The attributes and their levels are shown in Table 1.

Table 4.1 Attributes and levels in Börger et al. (2020)

Attributes	Levels
Seabed condition	<i>35%/50%/70%</i> area of seabed in the East Marine Plan restored to a natural condition
Coastal communities	<i>10%/25%/50%</i> of jobs in traditional industries lost in 10 years
Seabird population	<i>66%/40%/75%</i> of species in the East Marine Plan area with a healthy population
Additional tax (household per year)	<i>£0; £10; £20; £50; £75; £100; £150</i>

Note: Business-as-usual levels in italics.

Different elicitation formats can be used. “Pick one” asks respondents to choose the preferred alternative, "rating" to assign a score to each alternative (usually from 0 to 10), "rank order" to order the alternatives, or choose the most preferred and the least preferred alternatives. This question format is known as the Best-Worst method (Louviere et al., 2013). In Börger et al. (2020), each respondent is asked to complete nine choice cards and choose the most preferred alternative in each of the choice sets.

For a more comprehensive illustration of the implementation phases of a stated preference study and biases, refer to widely accepted guidelines (Riera et al., 2012; Johnston et al., 2017).

We use two examples to show the basic steps of conducting an SP survey, see Table 1. Specifically, Ramajo-Hernandez and Saz-Salazar (2012) will be used to illustrate how surveys are designed in a CV study. Börger et al. (2020) will be used to show the survey design in DCE methods. Ramajo-Hernandez and Saz-Salazar (2012) employ the contingent valuation method to evaluate the non-market benefits of water quality improvements in the Guadiana River Basin (GRB) in relation to the implementation of the European Water Framework Directive (WFD). Börger et al. (2020) use a discrete choice experiment to value a set of ecosystem services in the UK section of the North Sea, with a focus on offshore seabed habitats, regarding both use value and non-use values.

Table 1 - Illustrating the basic steps of conducting SP studies with examples

Steps in designing a stated preference survey	Contingent Valuation (CV) Ramajo-Hernandez and Saz-Salazar (2012)	Discrete Choice Experiments (DCE) Börger et al. (2020)
1. Identify the change(s) in quantity or quality to be valued	Water quality changes from the current state to the target, defined by chemical variables such as biological oxygen demand (BOD5) and the level of phosphorus (P).	The changes in three attributes (Seabed condition, Coastal communities, Seabird population, see Table 1). The changes are based on the Good Environmental Status, a management target under the Marine Strategy Framework Directive.
2. Identify the affected or reference population	The target population consists of those living in the Guadiana River Basin (GRB) area.	English population
3. Select a data collection mode	Face-to-face interview	A market research company conducted an online survey.
4. Choose a sample size	Based on population distribution, a stratified random sampling procedure was adopted. A total of 505 individuals living in 39 municipalities of the GRB area were included in this study, corresponding to 4.5% of the reference population.	Respondents were recruited from a pre-existing survey panel owned by Snap Surveys, using quota sampling to match the characteristics of gender, age group, and region. A total of 1,212 completed questionnaires were received.
5. Design the information component of the survey instrument	<p>Two types of visual aids are employed to make the CVM scenario more understandable. The first uses images that describe the current state of water quality and the target to be reached once the WFD (Water Framework Directive) measures are adopted. The second uses a water quality ladder (defined in terms of its suitability for specific recreational activities, ranging from “bad” to “excellent”) and asks respondents to value the changes in water quality from the current status to the target quality status, “very good”.</p> <p>The payment vehicle is an increase in the water bill, as it is believed to be plausible and familiar to the surveyed population, thereby adding credibility to the hypothetical market.</p> <p>The time frame of the payment is every time the water bill is paid.</p>	<p>The selection of attributes to be valued was based on a review of marine valuation studies in the UK and further refined in an international expert and stakeholder workshop.</p> <p>The development of attributes and their descriptions, as well as the entire questionnaire, was conducted by an interdisciplinary research team and based on recent recommendations for stated preference methods.</p> <p>Focus groups were held in four cities (Plymouth, Birmingham, Hull, and Ipswich) across England to assess attributes and levels, as well as the questionnaire content. Nine participants (selected to ensure a balance in terms of age, income, and gender) attended each focus group.</p> <p>The payment vehicle is a tax. The tax must be paid by all households in England every year for the next five years. The plan will be realized if the willingness to pay covers the cost of implementing it.</p>
6. Design the valuation question	<p>A dichotomous-choice question was used. Ten different bids were used, ranging from a minimum value of 0.5 to a maximum of 9.0 euros, based on the responses from a pilot study.</p> <p>They allowed for values of zero, asking respondents before the dichotomous-choice question if they were willing to contribute financially to the project by increasing their water bill.</p> <p>To the respondents who stated they were not willing to pay, a follow-up question was asked to differentiate protest responses from true zero responses.</p>	<p>A pick-one question format was used.</p> <p>The business-as-usual options, offered at a zero price, allowed the respondents to assign a zero value.</p>
7. Develop auxiliary questions for statistical analyses of valuation responses	They also collect information on attitudinal questions (about the environment in general and specifically about water resources and their pollution) and socio-demographic questions.	They collect information on whether respondents have visited the EMP area and the types of activities they engaged in there, as well as on UK coasts more generally. A set of attitudinal and socio-demographic questions is also included.
8. Pre-test and implement the survey	The pre-testing procedure consists of two focus groups and a pilot study of 50 interviews, aiming to identify sources of bias and clarify any unclear wording.	A pilot survey with 92 participants was conducted.
9. Analyze data	Analysis	Analysis

References

- Alfnes F., Yue C., Jensen H.H. (2010). Cognitive dissonance as a means of reducing hypothetical bias. *European Review of Agricultural Economics*, 37(2), 147–163.
- Bishop R.C., Heberlein T.A. (1979). Measuring values of extra market goods: Are indirect measures biased? *American Journal of Agricultural Economics*, 61(5), 926-930.
- Börger, T., Hooper, T. L., Austen, M. C., Marcone, O., & Rendón, O. (2020). Using stated preference valuation in the offshore environment to support marine planning. *Journal of Environmental Management*, 265, 110520.
- Boyle K. (2017). Contingent Valuation in Practice. In Champ P.A., Boyle K., Brown T.C. (eds), *A Primer on Non-market Valuation* (pp.83–131). Dordrecht: Springer.
- Carson R.T., Groves T. (2007). Incentive and informational properties preference questions. *Environmental and Resource Economics* 37:181–210.
- Caussade S., Ortúzar J.D., Rizzi L.I., Hensher D.A. (2005). Assessing the Influence of Design Dimensions on Stated Choice Experiment Estimates. *Transportation Research Part B*, 39: 621–640.
- Champ P. A., Moore R., Bishop R. C. (2009). A comparison of approaches to mitigate hypothetical bias. *Agricultural and Resource Economics Review*, 38(2): 166–180.
- Cummings R.G., Taylor L.O. (1999). Unbiased value estimates for 523 environmental goods: a cheap talk design for the contingent valuation method. *American Economic Review*, 89:649-65.
- Davis R.K. (1963). Recreation Planning as an Economic Problem. *Natural Resources Journal*, 3(2), 239–249.
- Haab T., McConnell K. (2002). *Valuing environmental and natural resources: the econometrics of non-market valuation*, Edward Elgar, Cheltenham.
- Hammack J., Brown G. (1974). *Waterfowl and Wetlands: Toward Bioeconomic Analysis*. Baltimore: Johns Hopkins University Press.
- Hensher DA, Rose J, Greene W. (2015). *Applied Choice Analysis. Second Edition*. Cambridge: Cambridge University Press
- Johnston R.J., Boyle K.J., Adamowicz W.V., Bennett J., Brouwer R., Cameron T. A., Hanemann W.M., Hanley N., Ryan M., Scarpa R., Tourangeau R., Vossler C. A. (2017). Contemporary Guidance for Stated Preference Studies. *Journal of the Association of Environmental and Resource Economists*, 4 (2), 319–405.
- Ladenburg J., Olsen S.B. (2014). Augmenting short cheap talk scripts with a repeated opt-out reminder in choice experiment surveys. *Resource and Energy Economics*, 37, 39-63.
- Louviere J., Lings I., Islam T., Gudergan S., Flynn, T. (2013). An introduction to the application of (case 1) best-worst scaling in marketing research. *International Journal of Research in Marketing*, 30, 292–303.
- Mariel, P.; Hoyos, D.; Meyerhoff, J.; Czajkowski, M.; Dekker, T.; Glenk, K.; Jacobsen, J.B.; Liebe, U.; Olsen, S.B.; Sagebiel, J.; et al. (2021). *Environmental Valuation with*

Discrete Choice Experiments: Guidance on Design, Implementation and Data Analysis, Springer Nature: Basingstoke, UK, p. 129.

Mitchell, R.C., Carson, R.T. (1989). Using Surveys to Value Public Goods: The Contingent Valuation Method. Resources for the Future, Washington DC.

Notaro S., De Salvo, M. (2010). "Estimating the economic benefits of the landscape function of ornamental trees in a sub-Mediterranean area", *Urban Forestry & Urban Greening*, 9 (2), 71-81.

Ramajo-Hernández, J., & del Saz-Salazar, S. (2012). Estimating the non-market benefits of water quality improvement for a case study in Spain: A contingent valuation approach. *Environmental Science & Policy*, 22, 47-59.

Riera P., Signorello G., Thiene M., Mahieu P.A., Navrud S., Kaval P., Rulleau B., Mavsar R., Madureira L., Meyerhoff J., Elsasser P., Notaro S., De Salvo M., Giergiczny M., Dragoi S. (2012). Non-market valuation of forest goods and services: Good practice guidelines. *Journal of Forest Economics*, 18(4), 259–270.

Rose J.M., Bliemer M.C.J. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews*, 29(5), 587-617.

Scarpa R., Notaro S., Louviere J., Raffaelli R. (2011). Exploring Scale Effects of Best/Worst Rank Ordered Choice Data to Estimate Benefits of Tourism in Alpine Grazing Commons, *American Journal of Agricultural Economics*, 93(3), 813-828.

Vossler C.A., Holladay J.S. (2018). Alternative value elicitation formats in contingent valuation: mechanism design and convergent validity. *Journal of Public Economics*, 165, 133–45.

Welsh M.P., Poe G.L. (1998). Elicitation effects in contingent valuation: comparisons to a multiple bounded discrete choice approach. *Journal of Environmental Economics and Management*, 36, 170–185.