Water Valuation and the Application of Declared Preferences Method: Contingent Valuation vs a Choice Experiment Approach

Óscar Alfranca
Polytechnic University of Catalonia

In the Declared Preferences methods (DP) questions are developed to different persons in order to establish alternative hypothetical scenarios. The most important methods which are based in a Declared Preferences hypothesis are the Contingent Valuation (CV) method and those which are based in a Choice Experiment (CE) Approach. In DP methods questions are used in order to establish preferences over hypothetical scenarios, which could be different (either goods or services). Each alternative is described through different attributes, and the answers can be used in order to determine the different preferences. Some very relevant methods which are based in the DP hypothesis can be found which are based in CV and CE methods.

Keywords: declared preferences, revealed preferences, experiments of election, choice experiments

INTRODUCTION

If the Declared Preferences (DP) method is valid for the estimation of some environmental values, then these methods could be useful for addressing problems such as, for example, what might be the consequences of holding a referendum on some environmental issue. Of special interest could be, for example, some non-use methods.

Some valuation methods are based on the DP hypothesis. The main methods are based on Contingent Valuation (CV) and Choice Experiments (CE), although another relevant method could be, for example, Contingent Behavior (CB). Questions related to CC focus on potential hypothetical behaviors (e.g., how willing would people be to visit a park if an entrance fee were imposed at a certain price).

Valuation methods based on DP techniques allow to design the approximation of the personal values of goods related to the environment. A principal virtue of these methods lies in their conceptual simplicity. Another important method that is based on DP is the CE. In this case, the CE are described by a finite number of attributes, which can refer to different levels, and thus a good is described by a limited number of characteristics, whose individual attributes can take different levels. A good can be characterized by a limited number of attributes (i.e., the specific alternatives of the different factors)\(^1\). Water valuation can be carried out by applying any of the DP methods, such as CV and CE, CB, among others. Questions based on the CB method raise questions such as “How many times could a park be visited if a 2 euro entrance fee were imposed?” In contrast to the CV questions, those related to the CE method focus on Hypothetical Behavior.

Among the objectives of some of the studies, conducted using the DP method, are, among others, the collection of data to maximize the validity and reliability of the estimated results. As is commonly reflected
in the literature, the DP-based valuation methods and the sample used should be effectively collected with
the desired properties.

In the application of an effective research method based on the DP, (with which to design the
instruments that could be used in the environmental valuation), and maintaining the status quo in the
sample. In this case we find the selection of a random sample, with a chosen population and the selection
of a method with the fundamental environmental conditions.

The conflict between the recommendations of the guidelines, the evidence in contemporary
environmental literature and the design of an environmental questionnaire, based on environmental
practices and based on the DP method, is of particular interest. The recommendations described here are
based on a review of the literature on DP methods, applied in different fields, such as environmental
economics, health economics or transport economics.

WATER VALUATION AND DP METHODS

In DP methods, questions are asked to individuals in order to establish their preferences about
alternative hypothetical scenarios (either goods or services). Each alternative is described through different
attributes, and the answers can be used in order to determine the different preferences. Some very relevant
methods which are based in the DP hypothesis can be found which are based in CV and CE methods.
Examples based on CV can be, in some cases, controversial. Although the CE method has been frequently
applied in the treatment of reclaimed water, this method has not been generally used in water management.

The CE is a quantitative technique based on the DP method (which allows the establishment of
individual particularities), which enables to discover how different individuals value various attributes of a
program, product or service by means of questions based on the establishment of a choice on hypothetical
alternatives. The CE are applied in a variety of methods, which can be related to business planning, resource
location or health policy, among others.

The application of the DP technique makes it possible to discover how individuals value the various
attributes of a good. The CE is a quantitative technique based on the application of DP. CE is a quantitative
method, which can be used in the absence of data on Revealed Preferences (RP). The method involves and
develops questions to ask individuals to establish their preferences about alternative hypothetical scenarios
(either goods or services). The CE have been applied in diverse methods that include aspects related to
health policy, business planning, resource location, or their relationship with the level of income. Although
the CE method has been frequently applied in the treatment of reclaimed water, this method has not been
generally used in water management.

APPLICATION OF THE DP METHOD

DP-based methods can provide estimates regarding the values associated with changes in welfare. The
status quo level makes it possible to propose relative changes, which are usually established to be credible,
and thus to answer possible questions. Thus, possible scenarios are presented, based on the DP hypothesis,
and in stages that allow the application of a method based on this same hypothesis, and linked to the
description of the scenario.

Description of the Scenario

The questionnaires should describe the type and extent of the scenario change, using language that is
tailored and understandable to the respondent. The description of this scenario also requires information
regarding the overall mechanism (usually a policy, investment or management review), which would
generate changes that could be assessed. The mechanism is described at a practical level, (so that it would
be consistent and real), which could induce change.

The questionnaires should describe the type and extent of the change, using language that is tailored
and understandable to the respondent. The description of the scenario also requires information about the
mechanism or set of mechanisms (typically a policy revision, investment or management change) that may
generate the business that should be assessed. In other words, the mechanisms should be described in a practical and consistent way, with a real world, which could be modified\textsuperscript{3}. Inaccurate and qualitative terms (such as “high”, “medium” and “low”) should be avoided until they are clearly defined by those giving the related answers. From the above, intermediate changes (or processes to generate changes), produced by those who give the answers, should also be valued\textsuperscript{4}.

Baseline conditions and their changes, as well as other scenario assessment components, may include subjective perceptions for the responses\textsuperscript{5}. The scenario design should take subjective perceptions into consideration when describing the change to be assessed. This would include the presentation of information in a way that people would understand and accept, and likewise, deliberation about strategies would be used in order to evaluate responses.

**Pre-Test Exposition**

The Pre-Test method is a comprehensive, specific method based on a guide and a questionnaire\textsuperscript{6}, based on the DP\textsuperscript{7}. Pre-test quality is a necessary component of the design, since it is essential to assess whether the number of interviews conducted is appropriate, and each of them may vary in context. For most applications, a minimum of four to six groups is recommended, and a higher number is suggested for new, unfamiliar, or difficult to quantify assets.

The tests are usually performed using data from pilot studies, which allow the design of DP studies, in which preliminary response rates and behaviors are established, with statistical hypothesis testing. Any type of trial could be conducted using population-based tests, whenever possible. The design of the study should reach its screening test, including the selection of the chosen groups or the conduct of interviews, which would allow the characteristics of the participants to be obtained, as well as the application of selection methods.

Pre-testing could be documented including types of pre-tests, numbers and characteristics of respondents, and sequences used to manage pre-tests. For quantitative pre-tests, the instruments, for data and surveys, should be retained, and their content should be maintained and documented. Audio and video recordings should be useful when the group’s substance or the interview content needs to be reviewed at a later date.

The record should be kept for key decisions, and for each step of the testing process. In addition, the results published and identified, as well as the decisions taken in the design phase, should also be represented.

**Approaches Based on the Presence or Absence of Water-Based Attributes**

The validity, reliability and applicability of DP studies depend on the explosion of the changes caused in the water. Even if there has been an increase in the use of CE in recent times, it is not clear whether CV or CE, in general, allow a higher valuation to be obtained. Each has advantages and disadvantages\textsuperscript{8}.

In contrast, CV offers opportunities to estimate values when an element cannot be easily defined in terms of attributes. The choice and differentiation between CV and CE is complex and should be based on the answers obtained and on the perceptions of the changes in values, the objective decision being considered, and the type of information required. For example, while the valuation application may be conceptualized and communicated as a set of attributes, responses to reports, respondents may or may not think of changes in terms of attributes.

On the other hand, CE may disclose information regarding the value of individual attributes, and their desire to modify some of them. These marginal values of individual attributes, which may remain invisible in a classical CV study.

The two approaches also have different advantages and disadvantages in terms of information provision (at least as commonly designated). Attributes presented using short descriptions, usually in CE, can sometimes simplify important features of the scenario\textsuperscript{9}.

Given such considerations, the choice of a question format should not preclude the attribution, or non-attribution, of methods based solely on factors such as ease of application or prevalence in the literature.
Some basic considerations are suggested when making decisions to apply the CV or CE format. First of all, the assessment of the specific characteristics of the articles (or the articles as a whole), and what the information needs of the decision-makers are, is modified. In some cases, such as the estimation of damage spills related to oil spills, the question can be linked to the estimation of specific values, and of a set of changes, as a whole.

CV allows for such supports in a decision-making context. In other applications, where a change may affect some attributes and not others, or where a range of different attribute changes need to be evaluated. In these cases, a CE may be more readily applicable because it provides marginal values for individual attributes of interest over a range of possible changes. This can be particularly useful when there is uncertainty about the physical and biological impact. In such cases, the valuation of a range of possible results by different attributes of levels in the CE provides a means of ensuring that the true impact has been assessed, or that values for a range of impacts are available to the decision maker.

Secondly, those who answer, think and value the changes, in terms of individual attributes or as a whole\textsuperscript{10}. One consideration is related to the DP structure, although respondents may think about the valuation of change in terms of attributes. The preference structure for these attributes could be more complex than simply linear or additive and separable preferences, which could involve the presentation of a matrix of CE questions and which are typically imposed as the assumptions of an econometric analysis on the response data.

In these cases, considerations would be offered if the CE models were sufficiently subtle to allow estimates of the ways in which different attributes might enter into the utility function of the respondents. In other words, if valid implicit prices could be obtained to represent (the willingness to pay for each attribute), and this attribute could be estimated.

The manner in which the presentation format influences, i.e., the understanding of what is to be valued, should also be considered. If the article is large and complex, it is difficult to describe how to adequately value the changes generated, and a set of attributes is used.

In this type of application, the use of attributes for communicating the changes to be valued can increase the scenario complexity, particularly when the number of attributes is large, or simply the number of attributes is not considered, or simple attribute descriptions are not applied\textsuperscript{11}. Complex choices may also trigger responses to cope with simplification, trial-and-error or response strategies inconsistent with decisions that would maximize utility\textsuperscript{12}. It is vital to keep in mind that these complexity-related challenges could affect the application of the CV method\textsuperscript{13}.

Third, how does the presentation of the information affect the understanding of the good to be valued? If the good is large and complex, it is likely to be difficult to describe how the change is valued adequately, using a parsimonious set of attributes\textsuperscript{14}. Complex choices may also trigger trial-and-error responses or response strategies that are not consistent with utility-maximizing decisions, and other methods\textsuperscript{15}, such as Systematic Conjectures\textsuperscript{16}, or Selective Inattention\textsuperscript{17} and Sequential Effects\textsuperscript{18}. The prevalence of anomalies may depend on several factors, including previous responses to knowledge or risks due to complexity\textsuperscript{19}.

**Experimental Design**

In this paper we discuss the Experimental Design (ED), in the context of a bidding in a CV exercise, discrete choice and an attribute-level CE assignment. ED in these contexts, and particularly for CE, is complex and elaborate\textsuperscript{20}.

The ED assumption defines the way in which different treatments are assigned to a question (or sequence of questions within a questionnaire)\textsuperscript{21}.  

4 Modern Sciences Journal Vol. 11(1) 2022
The primary objective of an ED, as applied to CV and CE methods, is to develop the means to obtain efficient and unbiased estimators, preference parameters and value estimators. Designs should make use of information obtained from empirical research, and require a pre-testing.

All treatment effects, and interactions with attributes, should be individually or jointly identified.

The ED should generally allow for interactions, (and perhaps other types of non-linear utility attributes), considering both statistical efficiency, cognitive capabilities and their estimates, constraints related to their implausibility, and the level of their attributes, as well as the combination in the design of the use, (which may be consistent in alternative model specifications), and considers how the levels related to the different attributes influence the design of their property.

Among the studies based on the DP method, a priori sources of information used in the formulation of the chosen designs, which can provide evidence to support sources of information, (which can be essential in their a priori formulation), should be distributed and used to design formulations, which can enable the design of evidence, and which provide evidence to enable the chosen design, and inform the different steps in the design of this process.

For a CV-related situation, investigators should choose the cost or quantity bid in consideration for the range and in relation to the space. The Effective Design for these questions ensures that the monetary quantity is credible to the respondents, it trains the estimated welfare in an unbiased way, and minimizes the variance of these estimates. Four components are involved in the basic design of the CE:

1. Selection of attributes, and the levels for each of them.
2. Decision on how many alternatives will be needed for each problem, as well as the number of questions for each subject.
3. Determination on how the levels of each attribute will be combined among the different alternatives.
4. Determination on how the attributes corresponding to the different questions are combined, can close the step for the presentation of the subsets of the respondents.

Combinations between questions and attributes can be randomly assigned to individual respondents, although this may result in a timely conclusion to increase the efficiency of the estimation.

Estimates of multi-attribute designs statistically enable a welfare that can be consistent with the model specification, and that can ensure the identification of important attribute-related as well as their interaction effects. The effective and statistically efficient multi-attribute design enables the estimated welfare that is consistent with the model specification, as well as ensuring the identification of the main effects and attributes of the interactive effects.

Experimental design decisions are informed by multiple factors, including change to be assessed, prior studies on specific applications, along with intuitions learned through surveys acquired through a design process. Attributes and levels should be selected and based on a combination of values needed to support implementation decision making, response plausibility and statistical efficiency. Levels should be chosen in light of the functional form to be used for the utility. For example, one might expect linear rates of change related to some utility, as well as its dimensions or non-linear changes, among others. The number of levels for each attribute should be sufficient to model these effects.

DP methods are often applied to inform public decisions, allocating knowledge sensitivity on issues related to ethics and research. Familiarity and compliance with recognized standards of ethical research, as well as the protection of human subjects, help to ensure that the stated preference is beyond doubt.

Data Collection and Ethics

DP methods are increasingly being used to inform public decisions, placing knowledge on issues related to ethics and research. The familiarity and conformity with norms or standards related to ethics research and the protection of human subjects can help to ensure that research on stated preferences is beyond doubt.

The DP protocol should be reviewed by universities or other boards to protect individuals, and informed consent should be obtained from them. The rights and welfare of human subjects (i.e., people) should be paramount, and their informed consent should be obtained from them. The study procedures should avoid deception, which could have significant negative consequences for respondents, unintended influence on
the results and validity of the study, or compromise the ability to use the study results to aid in decision making.

Mode Extension: Sampling and Non-Response Bias

DP research relies on multiple applications to collect data and sample data to identify respondents (including internet administration). Common modes include traditional email and in-person surveys, along with new approaches, such as the use of administration and the internet, mixed surveys with other surveys of the people being considered, mixed surveys and other electronic methods (such as tablets), and workshop evaluation. Each method has advantages and disadvantages, and there is no consensus regarding which method is the best. The use of personal interviews is the method recommended by the NOAA panel, and it allows interviews that can be followed in administrative processes, which could be subject to social desirability, and raise the related effects.

The most appropriate data collection is a specific context and the rationale for selecting how it should be documented. For all administration methods there are advantages and disadvantages. However, given the inability to effectively verbalize a complex evaluation of materials (such as a telephone conversation), this survey should be conducted with caution.

Samples should be collected from structures consistent with the populations whose values were estimated, and the respondents should be randomly selected from a structure in which an explicit procedure was used to select the sample. Contemporary approaches should be used in order to identify and mitigate the bias of those sample members who choose not to respond, including survey features and data collection to help identify and characterize the patterns of non-respondents. Wherever possible, the analysts could not rely solely on response rate and demographic data to infer the presence or absence of non-response bias.

Survey design and sampling considerations affect, and are affected by, multiple aspects of survey design, including sample structure and representativeness. A related survey requires a complete list of names/addresses to choose respondents. Those who respond on a website are often recruited via the Internet into a panel (and then a full survey is requested, once they have become panel members). Assuming the initial list, in which the samples are representative of the target population, these two approaches are response bottlenecks. For an e-mail survey, respondents can view the subject line and then decide whether to reply, introducing possibilities that the subject line is related to the same topic. For e-mail survey, the responders can look at the topic and then decide what form of response they will develop, introducing possibilities about the topic and bias.

For the panel survey and recruitment process, it may influence the behavior of the responses and the respondent, who should agree to become a panel that may not reflect the population through observable and unobservable characteristics. Through telephone administration, sample recruitment can be achieved with Random-Digit Dialing (RDD), but nowadays, due to the proliferation of cell phones, the use of dialing numbers has become more complicated, and the response rate has been reduced. Telephone lines established per household can vary based on income and other factors, affecting the likelihood that any household can be reached.

SURVEY MODE AND SAMPLING CONSIDERATIONS

These considerations affect and are affected by multiple aspects of the design, including the form of representativeness. In other words, a mail survey requires a complete list of names and addresses. A web-survey is often contracted to form a panel and then a full panel survey is asked once the members are part of the panel. Assuming the initial list from which the samples are selected and the target population, these two approaches present different impediments to response. For a summary e-mail, respondents can observe, and then decide, whether to replicate or not, introducing possibilities about selection and bias.

For conducting a web-based survey, the panel recruitment process may influence the panel from which those respondents are collected, and respondents who agree to be part of these panels may not reflect the population for all observable and unobservable characteristics. Using telephone administration, sample
recruitment can be achieved by RDD, but the proliferation of cell phones and number blocking has complicated sampling and reduced response rates. Telephone lines per household may vary according to income and other factors, which affect the likelihood that household members will be able to reach it.

The survey method and the way the survey is influenced by multiple design aspects, including representativeness. Recent research suggests that the mode of data collection does not substantially influence DP studies.

The market extension is usually defined, in principle, as a group of agents whose welfare would be influenced by the valuation of changes in a DP study. While the market extension could be difficult to identify for non-use values. In addition, differences between jurisdictions and locations of those who are affected may influence the understanding of the permissible market extension. For example, for many national policy decisions, residents of other countries affected by police typically do not hold up legally in a formal Cost-Benefit Analysis.

Determining the market extension is important to ensure that the correct market structure has been selected. It is also necessary when expanding (or aggregating), estimating population values (e.g., calculating the CBA). A clearly defined market area also makes it easier to determine whether the sample shape is available and complete (i.e., whether all affected individuals have a known probability of selecting the sample). A selected sample structure is essential for the identification of non-response effects.

**Survey Mode and Sampling**

Face-to-face interviews have the desirable feature that the completion process can be guided through trained interviews, but this approach is an expensive way to collect data and responses may be affected by unexpected or unanticipated interview-related effects. Telephone surveys are usually relatively inexpensive and are convenient for quick data collection.

Many countries provide access to representative sample structures, such as neighborhood address lists for email surveys.

**Response Rates and Non-Response Biases**

Most high quality surveys continue to be administered in an established manner as a method of increasing response rates. Many studies have found a weak relationship between response rates and non-response bias and the effort to increase the response rate.

**STIMULATION OF VALUE**

Methodological choices related to the stimulation of value can be classified into five categories:

1) Attribute and non-attribute based methods are applied.

2) The type of welfare measured to stimulate value.

3) The answer frame chosen to select the format of the chosen question.

4) The choice of payment method.

5) The use of ancillary questions and other design elements to support and assess validity.

The first of these is classified under the discussion of a survey, which has followed the proposed design. Categories 2-5 are listed here. Subsequently properties and incentives are discussed, along with other elements required to ensure reliable welfare estimation. The choice of approach depends on the contextual issue to be assessed, including the type of good, the nature of the decision being evaluated and the associated institutions (such as, for example, the existence of property rights or a tax system).

The literature shows that, in dealing with Public Goods, the ideal design includes choices, where the valued question is clearly understood, the payoff is complex if the proposed change is implemented, the respondent perceives his/her answers as influencing the provision of a thing that is being valued, and other aspects of the format’s stimulus (such as, for example, the number of questions and alternatives per question), incentivize true preference revelation.
Willingness to Pay vs. Willingness to Accept

There are multiple reasons to explain the possible differences between Willingness to Pay (WTP) and Willingness to Accept to Pay (WTA), along with the relationships between these empirical estimators and Hicksian welfare measures. From conceptual and theoretical perspectives of change under consideration at the same time, established empirical difficulties associated with WTA estimation lead most studies to estimate WTP. Some recent research attempts to address these empirical challenges, for example, via the use of stimulant mechanisms.

The decision about a general context determines whether WTP or WTA is the most appropriate measure for welfare, from a conceptual perspective, but the choice of welfare would be motivated by a combination of theoretical and empirical considerations, and the motivation for this choice should be explained. Since WTA estimation is often exposed to empirical challenges, such as difficulty in shaping compatible questions and increasing rejection rates of the WTP estimation scenario will lead to estimates with superior empirical properties, it may over-estimate or under-estimate true welfare measures if WTA is conceptually appropriate. However, the WTP estimate would not always be considered a default and the WTA estimate would be applied when appropriate and reliable.

In the work proposed by the NOAA panel, WTP estimation is recommended. However, with improvements since then, both in the understanding of why WTP and WTA should be differentiated and in the design of SP-based methods, this recommendation should not necessarily be satisfied.

In cases where payment reductions are institutionally feasible, questions of incentive compatibility may be designated. Examples include tax reductions associated with environmental changes.

The choice of WTP vs. WTA can be used and understood from the literature on this subject, supplemented by information from pre-testing. At the same time, established empirical difficulties associated with estimating WTA lead most studies to estimate WTP. WTA estimation can also be practical and appropriate.

The generalized decision context determines which of the two measures (either WTP or WTA) is the more suitable welfare measure from a conceptual perspective. The final welfare measure chosen, however, would be motivated by a combination of theory and empirical combinations, and the motivation for such a choice could be explained. Given the fact that estimation of WTA is often challenging, (such as difficulty in constituting compatible questions, and an increasing rate in a rejection scenario), estimation of WTP will often lead to estimates with superior empirical properties, but may overestimate or underpredict the true measure if WTA is conceptually appropriate. However, WTP estimation may not always be considered a default, and WTA estimation should be performed when appropriate and feasible.

The NOAA panel recommended the WTP estimate. In some situations where payment decreases are institutionally feasible, situations that represent incentives or payment decreases could and can be generated.

Response Question Evaluation Forms

There are multiple response formats for CV and CE methods. Questions in the form of responses and which configure contingent valuations include, but are not limited to, iterative bidding, open-ended extraction, payoff letters, and binary or dichotomous choice. These questions include several types of binary choice and multinomial format, along with several types of ranking, and which can refer to the best and worst type of good, although there are inconsistencies in the literature, as far as formats are concerned, under the estimation of the DP technique.

DATA ANALYSIS

Improvement in econometric data analysis is one of the main topics in the literature concerning DP methods. Improvements in econometric analysis are a prominent theme in the literature. There are many parametric, semi-parametric and non-parametric methods modeling alternatives. In particular, for a DP
function, the most appropriate estimators would be selected according to the data and the questions to be answered. Best practices for estimation and interpretation of models with similar data structure.

The recommendations focus on specific questions for DP data analysis or because it requires specific considerations. Advances in econometric analysis are a prominent theme in the literature. There are many parametric, semi-parametric and non-parametric alternatives related to modeling. Multiple classes of DP data are possible, and each poses a challenge that requires specific econometric methods. While econometric methods for analyzing some classes of DP data are virtually untested (e.g., inconclusive data), others have been subject to rapid development (e.g., discrete choice data). Current modeling often employs specific approximations to achieve a response pattern.

For example, the estimation often allows preference (and/or scale), heterogeneity and correlated responses with various types of conditional, mixed or generalized logit model. It should be noted that there is no particular model, or set of models, that can be recommended for all DP modeling. The selection of econometric estimators should reflect the specific aspects of the data to be analyzed, the hypotheses to be investigated and how the estimation will result in decision making. The exchange considered in the selection of the estimator should be explicitly documented. Modeling would be a selector by utility structure, by the theoretical structure assumed to motivate behavior in a DP context. Theory-utility, behavioral, statistical and other selection and specification models should be made explicit.

**PRINCIPAL DIFFERENCES BETWEEN CV AND CE**

Although the section emphasizes the differences between CV and CE, it is important to recognize that these approaches are close relatives in the Semi-Parametric valuation genre and could in principle be used to estimate equivalent values. These studies provide mixed results, but the large differences between the structure of the questions related to CV and the CE-type assessment are not always clear.

The difficulty of assimilating the effects of individual CV versus CE design elements, when comparing these approaches, speaks to the importance of collection formats as a whole. While the use of focus groups, cognitive interviews or other forms of pre-testing may help in the choice of an attribute, or the design of non-attributes.

**ESTIMATION OF VALIDITY**

Most of the literature published regarding methods for DP research focuses on the validity and reliability of the response value of the data. Although most research has shown that DP data are reliable. The main contention question has been validated, and the DP methods have been subjected to an extensive deployment of validity tests.

The prevalence of DP studies suggests a consensus that such analyses are important. As a precursor to our recommendations regarding this issue, this topic provides a brief introduction to the validity of the calculation on stated preference studies. Some individuals note a hypothetical bias about the validity of DP methods; of great interest have been DP studies, the value of which in some cases exceeds the value of the studies themselves. Most of the work has been conducted in the context of CV, but similar results have been found in CE work. The conclusion of a hypothetical bias is largely based on cash transactions in which an actual transaction (usually for a private good in an experimental setting), provides a benchmark against which DP estimators are compared. There is disagreement about the implications of these findings. Therefore, some authors emphasize that few of these investigations satisfy the compatibility requirements on the true revelation of preferences, and that it is not clear how experimental payoffs compare to a true payoff (value or criterion). Many valuation studies use an unimportant experimental treatment to obtain a behavior from which to compare cash transactions. There is an overwhelming evidence that this framework does not characterize response incentives for many responses, and validity tests lead in more than one representative field, toward biased experiments. In addition, research that is developed using voting-related behavior could generate a hypothetical bias, which is not universal.
STUDY REPORTS

Study reports are a critical element of the scientific method, enhancing the understanding of a paper and providing the basis for its replication. An effective communication is the cornerstone for future research in order to move forward with an empirical method and for the use and reuse of results from a given outcome, e.g., related to benefit transfer.

All studies, whether applied (to assist in decision making) or methodological (to assist in the evolution of research), should fully design the document, study design, implementation, analysis and results. Such transparency is essential for the scientific credibility of the studies and for their proper interpretation and utilization of the results. Documentation can also help to use the results for future decision making, even if it was not foreseen in the initial part of the study.

CONCLUSIONS

This paper describes the main steps for the application of a valuation process using the DP method. DP models can be combined with RP models, although this is only possible when they come from the same data-generating function. An area that has proven its conceptual importance in this literature is related to the examination of convergence or divergence in parameter estimation for discrete models.

This method is particularly useful in measuring the natural capital of ecosystem services. These services contribute to the measurement of human welfare, either through Direct or Indirect Methods.61

ENDNOTES

1. A recent example from among the CE can be found, for example, in Sarkar et al., 2019.
2. Some important examples of its use are those of the Exxon Valdez oil tanker, applied to assess the damage generated (Carson et al., 2003), as well as the activities related to the Kakadu National Park in Australia (Bennett, 1996). CV could provide, in some cases, a controversial view of environmental problems. This vision would be extended to various authors such as Hanemann, 1994. Bateman et al., 1997a; 1997b. Bateman, Munro & Poe, 2008. Carson, 2012.
3. From the above, the baseline and changes could be presented, in an adjusted, measurable and interpretable way, (Sarkar et al., 2012), and thus reflect results that would allow the assessment of utility and consequences that could be identified with those making the responses. Boyd and Krupnick, 2013. Johnston et al., 2016.
4. That is, these responses should be able to identify the links between intermediate changes and the final changes directly responsible for well-being, as proposed by authors such as Sarkar et al., 2012. Boyd y Krupnick, 2009; 2013. Boyd et al., 2016. Johnston et al., 2016.
5. For example, those that encourage a reaction from the adversary to avoid the rejection of a given scenario. Adamowicz et al., 1997. 2014. Cameron, DeShazo & Johnson, 2011.
6. A summary of the literature can be found in Presser et al., 2004.
8. For example, an attractive feature of CE is the ability to estimate marginal values, based on the attributes and the study design. However, this feature usually involves a disadvantageous trade-off, such as high complexity and potential loss with incentive compatibility, depending on the structure of the CE. Vossler, Doyon & Rondeau, 2012.
9. Conversely, the response may have difficulties identifying and distinguishing effects explained using narrative texts common in the CV method (Hoehn, Lupi & Kaplowitz, 2010). Both approaches, depending on the design, can present complex information.
10. For example, if the respondent thinks about a landscape or an ecosystem as a whole, then attributing this form could be inconsistent. In cases like this, the syntactic analysis of the totality as a whole could incorrectly characterize the way in which the components of the attributes or characteristics are understood or enhanced (Madureira, Nunes & Santos, 2005). On the other hand, individuals might think about changes, (e.g., in recreational locations), in terms of consistent traits, with a tabular, or attribute-based presentation (Hoehn, Lupi, & Kaplowitz, 2010).


This could lead to behavior based on simplification by discovery. Hoehn, Lupi & Kaplowitz, 2010.


16. As proposed by authors such as: Hensher & Rose 2009. Hensher & Layton 2010.


18. This method is considered to address potential anomalies that may arise in managing the response to simplicity levels. Day et al., 2012.

19. Some literature examples about the importance of complexity in the CV method can be found in authors such as: Sandorf, Campbell & Hanley, 2017. Hoehn, Lupi & Kaplowitz, 2010.


21. The ED assumption defines the way in which different treatments are assigned to a question (or sequence of questions within a questionnaire). The ED could represent, for example, how some blocks of questions are assigned to different versions of a survey (as could be, for example, bidding different levels of attributes). ED could be as simple as developing the supply structure for a simple, binary choice question, or as complex as a choice matter with multiple attributes, multiple alternatives within a question of choice, with multiple attributes, multiple alternatives within a question, multiple questions within a questionnaire, and multiple versions of survey instruments. Caussade et al., 2005.

22. The ED assumption defines the way in which different treatments are assigned to a question (or sequence of questions within a questionnaire). The ED could represent, for example, how some blocks of questions are assigned to different versions of a survey (as could be, for example, bidding different levels of attributes). ED could be as simple as developing the supply structure for a simple, binary choice question, or as complex as a choice matter with multiple attributes, multiple alternatives within a question of choice, with multiple attributes, multiple alternatives within a question, multiple questions within a questionnaire, and multiple versions of survey instruments. Caussade et al., 2005.


24. For a general discussion, one could follow, for example: Dillman et al., 2014.


27. As suggested by MacMillan et al., 2002.

28. As proposed by, for example: Johnson et al., 2013.

29. In this case, the compilation of methods that were not known at the time of the NOAA panel should be noted. Arrow et al., 1993.


31. These individuals are therefore excluded from both the market and the sample structure. Loomis, 1996.

Despite these characteristics, though they have experienced declining response rates (Brick & Williams, 2013) and have been forced for counting an increasing proportion of the population that could be reached via cell phones (e.g., Blumberg et al., 2012). In addition, cell phone numbers may not be immediately assigned to specific geographic areas, which would make matching responses to neighbor characteristics more complex.
32. However, the problem of random sampling of individuals at random among neighbors remains a challenge. Link et al., 2008. The Internet adoption has been one of the major recent developments for the gathering of data collections. Tourangeau, Conrad & Cooper, 2013.

33. Although in some cases, this indicator may be considered a poor indicator of non-response bias. Groves, 2006. Tourangeau & Plewes, 2013.

34. Some examples can be found in: Curtin, Presser & Singer, 2000. This impulse causes in some cases the worsening of the non-response bias, and at the same time the effort to improve the response rate, which in some cases worsens the non-response bias. Peytchev, 2009. Lundquist & Särndal, 2013.

35. Younjun Kim, Catherine L. Kling, and Jinhua Zhao Kim, review the literature concerning the differences between estimators explaining WTP and WTA, considering theoretical and empirical explanations. Kim, Kling & Zhao, 2015. This choice of welfare measures does not necessarily refer to legal subsidies, as indicated by Knetsch, 2007.

36. This logical reference condition of the response, as indicated by: Carson, Flores, & Meade, 2001. Kim, Kling & Zhao, 2015. As well as the use of financial incentives, to modify procedures related to behavior and health, Promberger, Dolan & Marteau, 2012. A proposal in this sense can be found in Johnston, Swallow & Weaver, 1999.

37. Notable examples include tax payments with environmental management (Johnston, Swallow & Weaver, 1999.), or the relationship with health (such as Promberger, Dolan, & Marteau, 2012.).


A discussion regarding this topic can be found in Kling, Phaneuf & Zhao, 2012, and Carson, Groves & List, 2014.

As indicated by authors such as: Carson, Groves & List, 2014.


This is the case of authors such as Costanza, D’Arge, Groot, Farber, among others, who estimate the Economic Value of 17 ecosystem services for 16 biomasses, based on published studies and a few original calculations. In order to calculate the full value of the biosphere, (most of which is obtained with results gathered outside the market), the estimated value is calculated at around $16-54 trillion per year, with an average of $33 trillion per year. Costanza et al., 1997.

REFERENCES


