Revista Mexicana de Ciencias Forestales Vol. 11 (59)
Mayo – Junio (2020)

DOI: https://doi.org/10.29298/rmcf.v11i59.676

Review article

Consideraciones básicas para la aplicación de experimentos de elección discreta: una revisión

Basic considerations for the application of discrete choice experiments: a review

Enrique Melo Guerrero[1], Rodrigo Rodríguez Laguna[1*], Miguel Ángel Martínez Damián[2], Juan Hernández Ortíz[3] y Ramón Razo Zárate[1]

Resumen

La necesidad de asignar un valor económico a los bienes y servicios ambientales derivó en el desarrollo de metodologías como los experimentos de elección discreta, que cobraron gran relevancia en los últimos años. La presente revisión se enfoca en mostrar los aspectos básicos por considerar para el uso empírico de este método en su proceso completo. Se incluyen el diseño experimental, la elaboración y aplicación de un cuestionario y una encuesta, análisis econométrico y las medidas de bienestar. Para ello, se hace una revisión bibliografía sobre los avances y aplicaciones de experimentos de elección en los últimos años; asimismo, se abordan y discuten los hallazgos sobre su confiabilidad y validez, así como los alcances y limitaciones de este método de preferencias declaradas. A pesar de los retos que enfrenta; por ejemplo, la complejidad de la tarea de elección y esfuerzo cognitivo por parte de los entrevistados, situación más evidente cuando se emplea en países en vías de desarrollo, es hoy en día el procedimiento preferido por los investigadores de la economía ambiental; ya que permite la descomposición del valor total del bien o servicio en el de sus atributos, y proporciona un enfoque más útil para la valoración económica, lo que resulta en mayores ventajas para los tomadores de decisiones en política ambiental.

Palabras clave: Conjuntos de elección, economía ambiental, medidas de bienestar, servicios ambientales, valoración económica, validez.

Abstract

The need to allocate an economic value to environmental goods and services resulted in the development of methodologies such as discrete choice experiments, which have gained great relevance in recent years. This review focuses on showing the basic aspects to be considered in the empirical application of the method throughout its entire process, which covers the experimental design, preparation and application of the questionnaire and survey, the economic analysis and welfare measures. For this purpose, the literature on the advances and applications of choice experiments in recent years was reviewed; specifically, we address and discuss the findings on reliability and validity, as well as the scope and limitations of this declared preferences method. Despite the challenges that it faces (complexity of the task of choice and cognitive effort on the part of the interviewees, a situation that is more evident in its applications in developing countries), it is the method preferred by environmental economics researchers so far, since it allows to break down the total value of the commodity or service into the value of its attributes. Furthermore, it provides a more focused approach that is useful for economic assessment, as well as greater advantages for environmental policy decision makers.

Keywords: Choice sets, environmental economics, welfare measures, environmental services, economic valuation, validity.

Fecha de recepción/Reception date: 27 de noviembre de 2019
Fecha de aceptación/Acceptance date: 20 de enero de 2020

[1]Universidad Autónoma del Estado de Hidalgo. México
[2]Colegio de Postgraduados, Campus Montecillo. México
[3]Universidad Autónoma Chapingo. México
*Autor por correspondencia: rodris71@yahoo.com

Introduction

The failure of economic activities to assess the environmental goods and services and the resulting problems of deterioration of the planet has triggered the emergence of tendencies to develop methodologies aimed at the economic assessment of the environmental assets within the categories of public goods and common resources that have a value, but which through time have lacked a price as well as markets where they can be exchanged. Throughout this process, the contingent assessment (CA) and the discrete choice experiments (DCEs) stand out for their easy empirical application. They are known as declared preferences methods and they are based on surveys for inferring the willingness to pay or to accept payment for a modification in the environmental quality of a natural asset that takes place through a hypothetical market.

The DCEs were developed by Louviere and Hensher (1982) and by Louviere and Woodworth (1983), based on the progress made in the research on markets and transportation, which were subsequently applied in other areas. Adamowicz *et al*. (1994) were the first to use the DCEs within the context of environmental resources; in this respect, today there are numerous researches (Cerda, 2011; Rocamora *et al.*, 2014; Brower *et al.*, 2016). Thus, DCE is a declared preferences method routinely utilized for environmental assessment purposes involving the prioritization of natural resources management policies.

The DCE methodology entails the generation and analysis of choice data through the construction of a hypothetical market presented to the survey respondents. For this purpose, the experiment is designed by generating several choice sets or choice cards containing mutually exclusive hypothetical alternatives for their assessment by the interviewees before choosing their preferred option. The options consist of the characteristics or attributes of the good, which are defined by one or more levels. The choice sets include a constant alternative that denotes the current situation, known as *statu quo*, and, at least, two plans for improvement.

The choices of the individuals imply tactical compensations between the attributes that are present in the alternatives on each choice card or set. The price or cost is included as

one more attribute, which makes it possible to convert the marginal utility estimates into estimates of willingness to pay (MWTP) due to changes in the levels of attribute, which, when combined, can yield welfare measures (Hoyos, 2010). This allows breaking down the total value of the environmental good into the values of its attributes, providing a more helpful approach to economic assessment, since it positions it in an environmental management and policy perspective (Riera and Mogas, 2006).

Despite the importance of the DCE method in the assessment of non-market goods, certain studies question the validity and reliability of its results, due to the existing evidence of violations of the theoretical assumptions (Raquetonarivo *et al.*, 2016). The purpose of the present document is to show the basic considerations that must be taken into account by researchers interested in the empirical application of the DCE method, as well as its scope and limitations, in order to promote its use and reduce the sources of bias.

Theoretical basis of the DCEs

The DCEs are founded upon the theories of welfare, of Lancaster's theory of consumers' choice and Ben-Akiva's and Lerman's probabilistic choice theory (1985). It is assumed that the choices of the survey respondents reveal the preferences of the individuals, who maximize their utility by providing a given price and budget level.

According to Hanemann (1984), the decision of the consumer can be divided into discrete and continual choices; the former, in relation to the good to be chosen, and the latter, in terms of the quantity in which the good is to be consumed. Tudela-Mamani and Leos-Rodríguez (2017) point out that Lancaster broke up with the traditional theory of the consumer's behavior when assuming that the consumers demand goods by virtue of their characteristics, which generate a given level of utility; therefore, the random utility theory assumes that a rational individual will chose whatever implies the highest expected utility.

Econometrics for the analysis of the DCEs

The analysis of the choices in the DCEs is based on the random utility approach developed by McFadden (1974), which creates a bond between the deterministic model and a statistical model of human behavior (Hoyos, 2010).

The specification of the econometric model requires the determination of two fundamental aspects: the definition of the utility function (the form in which the random term is included in the indirect utility function) and the distribution assumption regarding the error component. The most common definition of the utility function assumes that the error component (ϵ_{ij}) is entered as an additive term:

$$U_{ij} = V_{ij}\,(Z_{ij}\,S_i\,M_j) + \varepsilon_{ij}$$

Where:

U_{ij} = Utility function

V_{ij} = Deterministic component of the indirect utility function for each alternative *j* of the choice set *C*

Z_{ij} = Attributes

S_i = Socioeconomic characteristics of the users

$M_{i\,=}$ Income

The user *i* prefers the alternative *m* over any of the options of the choice set *C*, if and only if the utility of *m* is superior to that offered by each of the other options, *i.e.*, if $U_{im} > U_{ij}\ \forall\ m \neq j$, where *m* and $j \in C$. The probability of choosing alternative *m* is expressed as follows:

$$Pr(im)=Pr\,[(U_{im}) > U_{ij})\ \forall\ j \neq m] = Pr\,[(V_{im} - V_{ij}) > (\varepsilon_{ij} - \varepsilon_{im})] \qquad (1)$$

The observable component (indirect utility function, V_{ij}) can be expressed as a linear function of the explanatory variables (Blamey *et al*., 1999):

$$V_{ij}=\alpha_j+\beta_1 Z_1+\beta_2 Z_2+\ldots\beta_k Z_k+\gamma(M_i-Cost_j)+\delta_1(S_1*\alpha_j)+\ldots\ \delta_p(S_1*\alpha_j) \qquad (2)$$

Where

α_j = Alternative Specific Constant (ASC) for alternative *j*

β_k = Utility coefficients vector associated with explanatory variables vector Z_k

γ = Coefficient associated to the individual's income minus the cost of alternative *j, (M_i–$Cost_j$)*

δ_p = Coefficients vector of the socioeconomic variables

The ASC is the parameter for each particular alternative represents the role of the unobserved utility sources. According to Riera and Mogas (2006), when this element is omitted from the utility function in the estimation process, significant differences in the obtained welfare measures are produced. Although certain studies have opted for excluding the ASC, today the literature that includes it in the discrete choice analysis is more prevalent (Hoyos, 2010; Johnston *et al.,* 2017).

On the other hand, the error terms are assumed to be independent and identically distributed, with a Gumbel distribution or a Type I extreme value distribution; therefore, the choice model is estimated by means of a logit multinomial model (LMN) (McFadden, 1974; Louviere *et al*., 2000):

$$Pr(im)=\frac{exp^{\omega v_{ij}(Z_{ij},S_i,M_i)}}{\sum_j exp^{\omega v_{ij}(Z_{ij},S_i,M_i)}} \qquad (3)$$

The previous expression considers the attributes to be assessed, as well as the socioeconomic characteristics of the individuals; it contains a non-estimable, independent scale parameter of the function (Álvarez-Farizo *et al.*, 2005). The parameters of the indirect utility function in the LMN (α, β and δ) are estimated using the maximum likelihood method (Greene, 2003).

The main limitation of the LMN models is the implicit assumption of Independence of Irrelevant Alternatives (IIA), according to which the disturbances are independent and homoscedastic; *i.e.* the choice probability quotient of two alternatives is independent of any other real or potential alternative. Breach of this assumption leads to biased results (Louviere *et al.*, 2000). Therefore, the Hausman and McFadden test must be applied (Greene, 2003). However, the IIA hypothesis is not always sustainable, and therefore it eventually incorporates the correlation between alternatives (Hoyos, 2010).

In this regard, Train (2009) points out that the scope and limitations of the LMN models are: i) they can represent systematic variations related with the observed characteristics of the survey respondents (income level, schooling, family size), known as variations in taste, but they cannot represent random variations (preferences) of the survey respondent, known as taste variations; ii) compliance with the IIA assumption in the LMN models does not always reflect realistic situations; and iii) the LMN models handle situations in which unobserved factors are independent, but not when these generate a correlation through time, for example, between alternatives. This limitation is associated with market or labor studies in which the interviewees often participate in various choices through time, and their current decision may be influenced by that of the previous exercise.

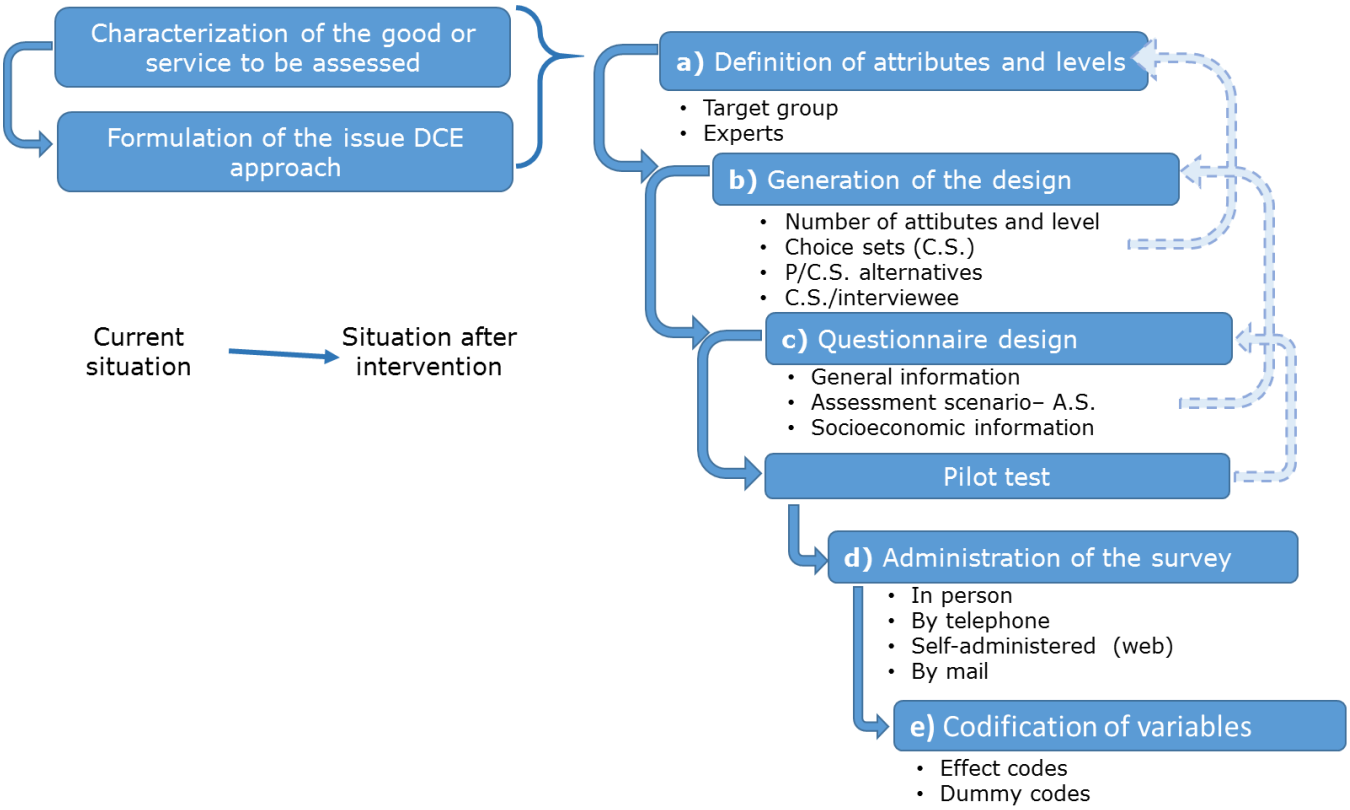
There are other generalized extreme value models that relax the IIA assumption, such as the nested logit model, the cross-nested logit model, the network model, the paired combinatorial logit model, and the generalized nested logit model; however, these may exhibit rigidity in their application, and they may have greater computational requirements.

The mixed logit (LMX) model is more flexible (McFadden and Train, 2000) and overcomes the limitations of the previously cited models; Tudela-Mamani and Leos-Rodríguez (2018) document that the LMX model "evades the limitations of the LMN model, allowing a random variation of preferences, unrestricted substitution patterns, and correlation between unobserved factors through time".

The use of an LMX model entails three specification problems (Hoyos, 2010): i) the determination of the parameters that must be modeled as randomly distributed; ii) the choice of the mixture distribution for the random coefficients, and iii) the economic interpretation of the randomly distributed coefficients. According to the empirical evidence that exists in the literature about the DCEs, the most commonly used models correspond to the LMN and LMX models (McFadden and Train, 2000; Hoyos, 2010; Espinal *et al*., 2014, Tudela-Mamani and Leos-Rodríguez, 2018).

Choice Experiment Design

The DCE procedure involves the following stages: A) the design of the DCE and B) estimation procedure and determination of the MWTP. This section addresses the DCE design phase, which in turn includes: a) definition of attributes and supply levels; b) generation of the experimental design; c) development of the questionnaire; d) administration of the survey, and e) codification of variables. Each of these must be carried out comprehensively and will allow providing sequential feedback to the improvement areas identified in the final design of the DCE. Figure 1 summarizes the design phase of the choice experiment.



Source: Prepared by the authors upon the basis of the DCE procedure

Figure 1. Choice experiment design.

Definition of attributes and levels

Before developing the experimental design, the good or service to be assessed must be characterized; for this purpose, the issue must be formulated in terms of policies or of a program for improvement that will reflect the change between the current situation and the situation expected to result from the intervention (Tudela, 2010). Therefore, several strategies may be used to obtain the information, including bibliographical review, participative diagnosis, interviews with key actors, forming discussion groups, among others.

Based on the attributes, the current situation (*statu quo*) of the forest service or good to be assessed is described to the interviewees; the levels represent the degree of intervention in which each of these will modify the current status. The attributes may be quantitative (reforestation percentage, level of water supply, reduction of the

erosion, etcetera), qualitative (aspect of the water, type of tourism, resilience of the ecosystem, etcetera), generic (the same levels are used for all the alternatives), or alternative-specific.

The process of identifying attributes and levels requires piloting strategies and discussion among target groups in order to ensure that the choice experiment design will incorporate the most important attributes and, thus, prevent uncertainty as to whether any relevant one was left out.

It must be kept in mind that a larger number of attributes and levels increases the quantity of both the choice sets and the alternatives that constitute them. According to Hoyos (2010), this situation may violate the assumption of independence from irrelevant alternatives and may have an impact on the conditions of regularity in the choices of the interviewees, thereby limiting the interval of applicability of discrete choice models. For example, adding an alternative that is similar to another within a set of options may increase the likelihood of choosing a third alternative merely because it is different (Hoyos, 2010).

The allocation of the levels of the attribute within the experimental design has an impact on the statistical power of the experiment in terms of its ability to detect potential statistical relationships between the data (Rose and Bliemer, 2014), which is directly related to the size of the study sample. However, the most common sample sizes in empirical researchers may compromise the ability to recover statistically significant parameter estimations, if a deficient design is selected. According to Breffle (2008), "although a large sample may compensate a deficient design, this is not necessarily the best way to approach the issue"; furthermore, this type o samples entail greater costs of the research.

When applying the DCEs empirically, it is possible to put forth more than one design, and there is a risk of not selecting the optimal one; therefore, it is crucial that the researcher determines precisely the functional form of the statistical model or utility function (Breffle, 2008; Johnston *et al.,* 2017). For this reason, it is important to characterize the issue accurately, depending on the techniques utilized to obtain the

information. Alpizar *et al.* (2001) mention that it is possible to approximate a utility function *ex-ante*, based on the results of previous studies on the same topic, on the judgment of experts, and on pilot tests, or else, to develop the sequential design approach proposed by Scarpa *et al.* (2007), in which the parameters are updated as more information is added.

Another important aspect regarding which decisions must be made is whether or not to include the *statu quo* as an attribute level. From the point of view of the consumer theory, this option needs to be considered (Hoyos, 2010); although it will reduce the efficiency of the design, since both the attributes and the levels will be increased. In this regard, the empirical evidence does not exhibit a consensus; while several authors include it in their studies (Espinal *et al.*, 2014; Rocamora *et al*., 2014; Villanueva *et al.*, 2017; Tudela-Manani and Leos-Rodríguez 2018), others omit it (Álvarez-Faricio, 2006; Birol *et al.,* 2006; Turner, 2013). Therefore, the inclusion or exclusion of the *statu quo* depends on the purpose of the study and must be assessed by the research team.

Generation of the experimental design

The experimental design is the organization of attributes and levels that constitute the alternatives in the choice sets; their identification and efficiency are the main statistical aspects involved in their construction (Hoyos, 2010). Their identification consists in the effects that can be independently estimated; i.e. the specification of the indirect utility function. Their efficiency is associated to the precision of the estimations of the parameters (Turner, 2013).

At this point, decisions must be made regarding: i) the number of attributes and their levels; ii) the number of alternatives, and iii) the number of choice sets (Breffle, 2008).

Once the attributes and levels have been identified, the factorial design, which can be complete or fractioned, must be defined. The complete factorial design includes all the possible combinations of the attributes and their selected levels and independently estimates the main effects and of the effects of the interaction. If the attributes and levels are numerous, a complete factorial design is impossible to apply, and therefore,

fractioned designs are resorted to, which comprise a sample of those attributes and levels and can be orthogonal (therefore minimizing the correlation between these (Bennett and Adamowicz, 2001). Statistical packages including SPSS or R are utilized to obtain them.

Based on the number of choice sets obtained, the tasks or choice sets per interviewee are determined. According to Oehlmann *et al.* (2017), there is a greater probability that the interviewees will choose the *statu quo* when the number of choice tasks per individuals is increased, when the intervals between levels are broader, and when the alternatives are similar; these situations are associated to a larger number of attributes and levels. In order to reduce the choice tasks per interviewee, different versions of the questionnaire are usually made; therefore, the orthogonal design includes the number of versions (blocks) as an additional attribute.

Development of the survey instrument

One of the most vulnerable aspects in the declared preferences methods is the use of questionnaires to obtain information; for this reason, the design attends to the recommendations included in literature for contingent assessment (CA). For example, the scenario is presented from the start in such a way that will allow the interviewee to understand the effects of the hypothetical program under consideration; furthermore, validation questions are integrated in order to verify that the interviewee has understood and accepted the proposed scenario, among others that are applicable to DCEs.

The questionnaires include at least three sections: i) general information and presentation of the ecosystem good or service to be assessed; the purpose of this is to enable the interviewees to make informed decisions based on a homogeneous understanding of the issue and its terminology (Turner, 2013); ii) assessment scenario or choice question; this must represent an incentive and must motivate the interviewees to provide their true preferences, and iii) socioeconomic information. The choice format must recreate, to the furthest possible extent, the context of a real choice (Turner, 2013); therefore, the selected payment method must be consistent with the context of the good or service to be assessed.

In order to obtain good welfare measures, it is necessary to establish a point of comparison with the scenario of change or improvement in the supply of the assessed good, also known as baseline or *statu quo*, which represents the option of "taking no action" and must be perfectly well understood by the interviewee if the assessment is to be effective. For this purpose, the use of images illustrating the current situation and recreating the intervention scenarios is recommended, for example, using current tools, such as graphic design.

Although the DCE has some advantages over the CA, such as the reduction of sources of bias —by making the interviewees compensate between the various characteristics of the good—, the breakdown of the total value into the value of its attributes, the obtainment of more information by the interviewees, and a decreased incidence of ethical protests, among others. Its application involves a larger level of complexity, not only for the researcher, when designing the experiment, but also for the interviewees as they carry out the choice task; this is due to the cognitive load demanded by this method, which renders it the main source of bias (Raquetonarivo *et al.*, 2016; Svenningsen and Jacobsen, 2018), particularly when the exercise is not sufficiently clear for the interviewee.

Therefore, once the questionnarie has been designed, it is advised to apply it within the target group for validation; subsequently, pilot tests must be carried out until the instrument is as fine-tuned and clear as possible.

These tests must verify several factors understood by the interviewees of the questionnaire (especially in the assessment scenario and the choice task) (Meyerhoff and Glenk, 2015) and of the considered attributes and levels, as well as the duration and the time, among others (Hoyos, 2010). Likewise, Carson *et al.* (1994) recommend doing "warm-up" choice exercises together with the interviewee in order to reduce the cognitive load and facilitate the understanding of the task.

Application of the survey

Today it is possible to apply the surveys in various ways: in person, by telephone, self-applied through digital platforms, or in mixed schemes (Champ and Welsh, 2007). Turner (2013) points out that, while the in-person or telephone surveys are equivalent to those implemented via internet, the latter are much quicker to apply and ensure homogeneity in the formulation of the issue and in the assessment exercise. However, this statement must be taken with reserves, as the cognitive demand of the exercise may result in greater sources of bias. Thus, further research on the reliability and precision of web-based surveys is required.

In general, literature recommends applying the interviews in person, rather than by telephone, while telephone interviews are preferred to email surveys.

In any case, there are three fundamental aspects to be considered in DCE surveys: i) sampling bias and non-response, ii) integrity of the experimental design, and iii) administrative cost. The sampling bias and non-response are characteristic of any survey-based research (McFaden, *et al.,* 2005; Turner, 2013); therefore, sampling must be as random as possible in order to obtain good estimates of the population parameters. For further readings on the advantages and disadvantages of each form of application of the surveys, it is suggested to make a deeper review of the two authors cited above.

The experimental design is a particular concern in the case of DCE surveys (Turner, 2013), and its efficiency greatly determines the success of a reliable sample. The administrative cost is determined by the sampling strategy and the selected confidence level, as well as by the strategies and incentives implemented by the researcher.

Codification of the variables

Previous to the econometric analysis, the explanatory variables must be codified in order to measure their effects on the model; this can be done by codifying the effects or by using dummy codes. The codification of the effects is more widely used in empirical DCE studies, as it prevents correlation with the codes arranged at the baseline and minimizes colinearity of the estimation matrices used for calculating the interactions (Hoyos, 2010). The most usual way of coding the variables is by following the procedure developed by Holmes and Adamowicz (2003), which consists in allocating values of -1, 0, or 1 to the variable, according to the interviewee's choice.

Table 1 shows the hypothetical codification for an attribute with two intervention levels that differ from the *statu quo*. In this case, a variable is generated for each level of intervention: A1 and A2.

Table 1. Hypothetic codification of the variables using the effect codes.

Interviewee's choice	**Attribute A**	
	Variable A1	**Variable A2**
Level A1	1	0
Level A2	0	1
Statu Quo	-1	-1

Source = Adapted from Tudela-Mamani and Leos-Rodríguez (2018).

Econometric analysis of the DCEs

Goodness-of-fit measures

The goodness-of-fit of an estimated model describes how well it fits the set of data and reflects the discrepancy between the observed and the expected values in the model. It is measured based on the statistic known as Pseudo-R^2 (or McFadden's R^2), which takes on values of 0 to 1, where 0 indicates a null or very low goodness-of-fit of the model, and 1 signifies that the model provides a perfect prediction of each choice. For the statistical analysis of DCEs, the Pseudo-R^2 values corresponding to models with a high goodness-of-fit range between 0.2 and 0.4; this interval is equivalent to an R^2 of 0.70 – 0.90 for a least ordinary squares regression (Tudela, 2010).

Choosing the right model

The econometric analysis includes more than one model before selecting the right one for the data of interest, for which purpose the following aspects must be considered (Tudela, 2010): i) the coefficients of the independent variables must have the expected signs; ii) the coefficients of the independent variables must be significant at a given confidence level; iii) the (Akaike) information criteria must be low, and iv) the model must have a better goodness of fit in terms of MacFadden's R^2.

Welfare measures

The parameters estimated based on a linear model (2) can be interpreted as marginal effects of the attribute of the good to be assessed over the possibilities of choosing one of the plans. Therefore, the marginal willingness to pay (MWTP) by intervention area is the willingness to pay for a unitary change in each, while the rest remains constant. The MWTP in any of the analyzed attributes is estimated based on the quotient that results from dividing the estimated coefficient of each attribute (β) by the rate increase coefficient (γ) (Alpizar *et al.*, 2001):

$$MWTP = -\beta\alpha^{*}\gamma^{-1}$$

Reliability and validity of the DCEs

Although DCE is one of the most widely accepted economic assessment methods today, its reliability and validity —*i.e*. whether the method provides consistent results with different survey designs that may be used to measure the same values— are constantly questioned. Because the assessment exercise is based on a simulated scenario, a "hypothetical bias" may occur (Raquetonarivo *et al.*, 2016), *i.e*. the preferences expressed by the survey respondents may differ when facing real economic circumstances (Hausman, 2012).

For this reason, validity and reliability tests of the method are constantly researched (Hess and Daily, 2014; Lancsar and Swait, 2014; Raquetonarivo *et al.*, 2016).

The validity may be external (convergent) or internal (theoretical and content-related). The tests for the former entail comparisons using different instruments of a DCE survey, while the internal validity tests focus on the basic assumptions of the method.

Internal validity

The results of the DCEs are theoretically valid when the choices of the survey respondents are aligned with the assumptions of the standard choice theory (on which the method is based) (Raquetonarivo *et al.*, 2016), defined by four axioms of expected utility maximization (Mas-Colell *et al.*, 1995): i) continuity, ii) monotonous preferences, iii) transitivity, and iv) stability.

The empirical evidence of the internal validity of the DCEs is mixed; while Carlsson and Martinson (2001) register no evidence of violation of the assumptions of transitivity and stability, other authors cite inconsistencies in the choice preferences of the interviewees, i.e. violations of the second axiom (Deshazo and Fermo, 2002; Rocamora *et al.*, 2014; Raquetonarivo *et al.*, 2016). On the other hand, Day and Prades (2010) point out the persistence of certain ranking anomalies in the DCEs that are similar to those cited in the contingent assessment method, which entails violations of the assumption of stability.

The researches that include rationality tests for identifying inconsistent responses have gained importance in the empirical applications of DCEs. However, their elimination in order to render the assessment more reliable may induce bias in the sample size and reduce the strength of the estimated models. In this sense, Rocamora *et al.* (2014) propose an iterative process for the detection and correction of inconsistencies in the choices by the interviewees, in order to minimize the loss of observations.

According to Carlsson *et al.* (2010) and Weller *et al.*, (2014), validity studies focus on the continuity axiom. Likewise, Colombo *et al.* (2013) provide evidence of the fact that some survey respondents in DCEs do not take into account all the attributes when they make their choices, or are not willing to compensate between them; this suggests a violation of the axiom and reduces the validity and reliability of the method in terms of producing reliable estimations of the economic value.

Recently, Sandorf *et al.* (2017) indicate that the level of knowledge of the survey respondents regarding the environmental good to be assessed influences the degree to which they ignore the attributes; a greater knowledge increases the likelihood of considering the attributes to be assessed; they also reveal a lower likelihood that the "cost" attribute will be considered. Therefore, questions reflecting the degree of knowledge of the good or service to be assessed must be added, and the assessment exercise must be made very clear for the interviewees, in order to avoid the omission of attributes leading to conditional estimations of willingness to pay, particularly to lower ones.

External validity

External validity tests may be criterion-related or convergent. The former compare between the actual and the hypothetical behaviors. In general, the DCEs are less prone to hypothetical bias than the CAs (Carson *et al.*, 1994; Alpízar *et al*., 2001; Carlsson and Martinsson, 2001) because the choice scenario is closer to a real situation of the interviewee.

The convergent validity tests deal with the correspondence between the measures obtained with DCEs and those obtained with other methods; the most widely used tests

for comparing the DCEs results are (Raquetonarivo *et al*., 2016): i) revealed preferences (travel costs, hedonic prices or production function); ii) contingent assessment or complete contingent classification techniques (Christie and Azevedo, 2009), or iii) through other assessment methods that may or not be consistent with the random utility theory, like the analysis of multiple criteria or attribute classification exercise (Azevedo *et al.*, 2009).

Other theoretical validity tests refer to the sensitivity to the scope, which assumes that the survey responders would be willing to pay more for an important effect than for a subset of this effect (Carson and Czajkowski, 2014). Although, in general, DCEs exhibit less biases than CAs, the empirical evidence points out that the DCE is more sensitive to the scope (Foster and Mourato, 2003).

The preferences obtained in the DCEs are more susceptible to the anchoring or bias effects of the point of departure and more complex, which leads to a strategic behavior by the interviewees. However, the empirical evidence is still not conclusive. Thus, Bateman *et al.* (2009) document proofs of this bias, while Day and Prades (2010) cite the opposite.

Conclusions

The use of discrete choice experiments has acquired great relevance in the environmental assessment across the world; today, the empirical applications for assessing environmental services are frequent. The variants that are present in the reviewed literature translate into a constant search for ways to perfect the method.

Discrete choice experiments offer more advantages than the contingent assessment, as they reduce sources of bias and provide a broader approach to economic assessment. However, this method still faces considerable challenges, such as the complexity of the choice tasks, given that it requires a greater cognitive effort by the interviewees. This situation is even more complex in the least developed countries, where the empirical applications are still rare and scarcely known. The interviewees face the task of exchanging complex forest goods and services in countries where the

level of perception and sensitivity to environmental issues differs vastly from the importance attached to these in countries with a greater environmental awareness and higher levels of education and income.

Various researches are focused on identifying inconsistencies in the DCEs due to violations of the basic assumptions of the economic theory induced by the complexity of the exercise; however, most of them agree in that the deficiencies can be alleviated by describing in detail the design phase of the experiment, as the changes systematically affect the estimations of the parameters and the variations of the error terms that lead to biased estimations.

The pressure exerted by the growing population on the natural resources emphasizes the need to have increasingly accurate instruments and tools that may facilitate environmental policy decision making.

In Mexico, there are 182 natural protected areas where the empirical applications of choice experiments are still scarce; therefore, we recommend increasing the case studies and orienting the use of this method for the assessment of environmental services toward the implementation of management policies that will incorporate the opinion of the various actors involved in the planning of everyday exploitation activities, as well as toward the preservation and improvement of these protected areas.

Acknowledgements

The authors wish to express their gratitude to Conacyt for the financial support provided for their postdoctoral stay and to the *Universidad Autónoma del Estado de Hidalgo* for endorsing the project.

Conflict of interests

The authors declare no conflict of interests.

Contribution by author

Enrique Melo Guerrero: design of the study, development of the methodology and drafting of the manuscript; Rodrigo Rodríguez Laguna, Miguel Ángel Martínez Damián, Juan Hernández Ortiz and Ramón Razo Zárate: counseling, review and editing of the manuscript.

References

Adamowicz, W., J. Louviere and M. Williams. 1994. Combining revealed and stated preference methods for valuing environmental amenities. Journal of Environmental Economics and Management 26: 271–292. Doi:10.1006/jeem.1994.1017.

Alpizar, F., F. Carlsson and P. Martinsson. 2001. Using choice experiments for non-market valuation. Economic Issues Journal Articles 81(1): 83-110.

Álvarez-Farizo, B., J. M. Gil y B. J. Howard. 2005. Evaluación de impactos ambientales derivados de estrategias de restauración. Economía Agraria y Recursos Naturales 5(10): 19-39. Doi:10.7201/earn.2005.10.02.

Azevedo, C., J. R. Corrigan and J. Crooker. 2009. Testing for the internal consistency of choice experiments. *In*: Edelstein, A., B. and D. Hauppauge (eds.). Handbook of Environmental Research. Nova Science Publisher. New York, NY, USA. pp. 507-517.

Bateman, I. J., B. H. Day, A. P. Jones and S. Jude. 2009. Reducing gain-loss asymmetry: a virtual reality choice experiment valuing land use change. Journal of Environmental Economics and Management 58(1): 106-118. Doi:10.1016/j.jeem.2008.05.003.

Ben-Akiva, M. and S. Lerman. 1985. Discrete choice analysis: theory and application to travel demand. MIT Press. Cambridge, MA, USA. 390 p.

Bennett, J. and W. Adamowicz. 2001. Some Fundamentals of Environmental Choice Modelling. *In*: Bennett, J. and R. Blamey (eds.). The Choice Modelling Approach to Environmental Valuation. Edward Elgar. Cheltenham, UK. pp. 37-69.

Birol, E., K. Karousakis and P. Koundouri. 2006. Using a choice experiment to account for preference heterogeneity in wetland attributes. Ecological Economics 60: 145-156. Doi:10.1016/j.ecolecon.2006.06.002.

Blamey, R., J. Gordon and R. Chapman. 1999. Choice modelling: assessing the environmental values of water supply options. The Australian Journal of Agricultural 43(3): 337-357. Doi: 10.1111/1467-8489.00083.

Breffle, W. 2008. In Pursuit of the Optimal Design: A Guide for Choice Experiment Practitioners. *In:* Giraud-Cullen, K. (comps). Benefits and Costs of Resource policies Affecting Public and Private Lands. University of New Hampshire. Durham, NH, USA. pp. 1-19.

Brouwer, R., M. Bliem and M. Getzner. 2016. Valuation and transferability of the non-market benefits of river restoration in the international. Ecological Engineering 87:20-29. Doi:10.1016/j.ecoleng.2015.11.018.

Carson, R. T., J. Louviere, D. A. Anderson, P. Arabie, D. S. Bunch, D. Hensher, R. M. Johnson, W.F. Kuhfeld, D. Steinberg, J. Swait, H. Timmermans and J. B. Wiley. 1994. Experimental analysis of choice. Marketing Letters 5. Kluwer Academic Publishers. Dordrecht, The Netherlands. pp. 351–368. http://link.springer.com.conricyt.remotexs.co/content/pdf/10.1007/BF00999210.pdf (10 de junio de 2019).

Carlsson, F. and P. Martinsson. 2001. Do hypothetical and actual marginal willingness to pay differ in choice experiments? Journal of Environmental Economics and Management 41: 179–192. Doi: 10.1006/jeem.2000.1138.

Carlsson, F., J. H. García and A. Lofgren. 2010. Conformity and the demand for environmental goods. Environmental and Resource Economics 47(3): 407-421. Doi:10.1007/s10640-010-9385-2.

Carson, R. T. and M. Czajkowski. 2014. The discrete choice experiment approach to MVC. *In*: Hess, S. and A. Daly. (eds.), Handbook of Choice Modelling. Edward Elgar Publishing. Northampton, MA, USA, pp.202-235.

Cerda, C. 2011. Aplicación de experimentos de elección para identificar preferencias locales por opciones de conservación y desarrollo. Bosque 32(3): 297-307. Doi:10.4067/S0717-92002011000300011.

Champ, P. A. and M. P. Welsh. 2007. Survey Methodologies for Stated Choice Studies. *In*: Kanninen, B. J. (ed.). Valuing environmental amenities using stated choice. Springer. Dordrecht, The Netherlands. pp. 21–42.

Christie, M. and C. Azevedo. 2009. Testing the consistency between standard CV, repeated contingent valuation and choice experiments. Journal of Agricultural Economics 60(1): 154-170. Doi: 10.1111/j.1477-9552.2008.00178.x.

Colombo, S., M. Christie and N. Hanley. 2013. What are the consequences of ignoring attributes in choice experiments? Implications for ecosystem service. Ecological Economics 96: 25-35. Doi:10.1016/j.ecolecon.2013.08.016.

Day, B. and J. L. Prades. 2010. Ordering anomalies in choice experiments. Journal of Environmental Economics and Management 59(3): 271–285. Doi: 10.1016/j.jeem.2010.03.001.

Deshazo, J. R. and G. Fermo. 2002. Designing choice sets for stated preference methods: the effects of complexity on choice consistency. Journal of Environmental Economics and Management 44(1): 123–143. Doi: 10.1006/jeem.2001.1199.

Espinal M., N. E., J. D. Gómez Z., A.D. Ramos R., Alzate T. y B. Mesa Z. 2014. Valoración económica del museo de arte Medellín. Aplicación de experimentos

de elección. Ensayos de Economía 24(45): 107-128. https://revistas.unal.edu.co/index.php/ede/article/view/50502 (23 de noviembre de 2018).

Foster, V. and S. Mourato. 2003. Elicitation format and sensitivity to scope. Environmental and Resource Economics 24(2): 141–160. Doi:10.1023/A:1022856329552.

Greene, W. 2003. Econometric Analysis. Pearson Education. New York, NY, USA. 188 p.

Hanemann, W. 1984. Discrete/continuous models of consumer demand. Econometrica 52(3): 541-561. Doi: 10.2307/1913464.

Hausman, J. 2012. Contingent valuation: from dubious to hopeless. Journal of Economic Perspectives 26(4): 43-56. Doi:10.1257/jep.26.4.43.

Hess, S. and A. Daly, 2014. Handbook of choice modelling. Edward Elgar Publishing. Cheltenham, UK. pp. 1-7.

Holmes, T. and W. Adamowicz. 2003. Atribute-Based Methods. *In*: Champ, P., J. K. Boyle and T. C. Brown (eds.). A primer Nonmarket Valuation. Kluwer Academic publishers. Dordrecht, The Netherlands. pp. 171-219.

Hoyos, D. 2010. The state of the art of environmental valuation with discrete choice experiments. Ecological Economics 1595-1603. Doi: 10.1016/j.ecolecon.2010.04.011.

Johnston, R. J., K. J. Boyle, W. Adamowicz, J. Bennett, R. Brouwer, T. A. Cameron, W. M. Hanemann, N. Hanley, M. Ryan, R. Scarpa, R. Tourangeau and C. A. Vossler. 2017. Contemporary Guidance for Stated. Journal of the Association of Environmental and Resource Economists 4(2): 319-405. Doi:10.1086/691697.

Lancsar, E. and J. Swait. 2014. Reconceptualising the external validity of discrete choice experiments. Pharmaco Economics 32 (10): 951-965. Doi:10.1007/s40273-014-0181-7.

Louviere, J. J. and G. Woodworth. 1983. Design and analysis of simulated consumer choice or allocation experiments: an approach. Journal of Marketing Research 20:350-367. Doi: 10.1177/002224378302000403.

Louviere, J. and D. A. Hensher. 1982. On the design and analysis of simulated choice or allocation experiments in travel choice modelling. Transportation Research Record 890: 11–17. http://onlinepubs.trb.org/Onlinepubs/trr/1982/890/890-003.pdf (30 de junio de 2019).

Louviere, J. J., D. Hensher and J. Swait. 2000. Stated choice methods: analysis and application. Cambridge University Press. Cambridge, UK. pp. 399.

Mas-Colell, A., M. Whinston and J. Green. 1995. Microeconomic Theory. Oxford University Press. New York, NY, USA. pp. 223.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behaviour. *In*: Zarembka, P. (ed.). Frontiers in econometrics. Academic Press. New York, NY, USA. pp. 105-142. https://eml.berkeley.edu/ reprints/mcfadden/zarembka.pdf (13 de junio de 2019).

McFadden, D. L., A. C. Bemmaor, F. G. Caro and J. Dominitz. 2005. Statistical Analysis of Choice Experiments and Surveys. Marketing Letters 16(3-4): 183-196. Doi:10.1007/s11002-005-5884-2.

McFadden, D. and K. Train. 2000. Mixed MNL models for discrete response. Journal of Applied Econometrics 15: 447-470. Doi: 10.1002/1099-1255(200009/10)15:5<447::AID-JAE570>3.0.CO;2-1.

Meyerhoff, J. and K. Glenk. 2015. Learning how to choose—effects of instructional. Resource and Energy Economics 41: 122-142. Doi:10.1016/j.reseneeco.2015.04.006.

Oehlmann, M., J. Meyerhoff, P. Mariel and P. Weller. 2017. Uncovering context-induced *Statu quo* effects in choice experiments. Journal of Environment Economics and Management 81:59-73. Doi:10.1016/j.jeem.2016.09.002.

Raquetonarivo, O., M. Schaafsma and N. Hockiel. 2016. A systematic review of the reliability and validity of discrete choice experiments in valuing non-market environmental goods. Journal of Environmental Management 183: 98-109. Doi: 10.1016/j.jenvman.2016.08.032.

Riera M., P. y J. Mogas. 2006. Una aplicación de los experimentos de elección a la valoración de la multifuncionalidad de los bosques. Interciencia 31(2): 110-115. https://www.redalyc.org/pdf/339/33911306.pdf (23 de noviembre de 2018).

Rocamora, B., S. Colombo y K. Glenk. 2014. El impacto de las respuestas inconsistentes en las medidas de bienestar estimadas con el método de experimentos de elección. Economía agraria y Recursos Naturales 14(2): 27-48. Doi:10.7201/earn.2014.02.02.

Rose, J. and M. Bliemer. 2014. Stated choice experimental design theory: the who, the what and the why. *In*: Hess, S. and A. Daly (ed.). Handbook of choice modelling. Edward Elgar Publishing. Cheltenham, UK. pp. 152-177.

Sandorf, E. D., D. Campbell and N. Hanley. 2017. Disentangling the influence of knowledge on attribute nonattendance. Journal of Choice Modelling 24: 36-50. Doi:10.1016/j.jocm.2016.09.003.

Scarpa, R., D. Campbell and W. G. Hutchinson. 2007. Benefit estimates for landscape improvements: sequential bayesian design rationality in a choice experiment. Land Economics 83(4):617-634. Doi:10.3368/le.83.4.617.

Svenningsen, L. S. and J. B. Jacobsen. 2018. Testing the effect of changes in elicitation format, payment vehicle. Journal of Choice Modelling 29:17-32. Doi: 10.1016/j.jocm.2018.08.001.

Train, K. 2009. Discrete choice methods with simulation. 2nd edition. Cambridge University Press. New York, NY, USA. pp. 338

Tudela, J. 2010. Experimentos de elección en la priorización de políticas de gestión en Áreas Naturales Protegidas. Desarrollo y Sociedad 66: 183-217. https://revistas.uniandes.edu.co/doi/pdf/10.13043/dys.66.6%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20%20 (23de noviembre de 2018).

Tudela-Mamani, J. W. and J. A. Leos-Rodríguez. 2018. Estimation of economic benefits due to improvements in basic sanitation services through choice experiments. Revista Chapingo Serie Ciencias Forestales y del Ambiente 24(2): 237-250. Doi:10.5154/r.rchscfa.2017.05.037.

Turner, R. W. 2013. Using contingent choice surveys to inform national park. Journal of Environmental Studies and Sciences 3(2): 120-138. Doi:10.1007/s13412-013-0110-7.

Villanueva, A., J. A. Gómez-Limón y M. Rodríguez-Entena. 2017. Valoración de la oferta de bienes públicos por parte de los sistemas agrarios: el caso del olivar de montaña en Andalucía. Economía Agraria y Recursos Naturales 17(1): 25-57. Doi:10.7201/earn.2017.01.02.

Weller, P., M. Oehlmann, P. Mariel and J. Meyerhoff. 2014. Stated and inferred attribute non-attendance in a designs approach. Journal of Choice Modelling 11: 43-56. Doi:10.1016/j.jocm.2014.04.002.

All the texts published by **Revista Mexicana de Ciencias Forestales –**with no exception– are distributed under a *Creative Commons* License Attribution-NonCommercial 4.0 International (CC BY-NC 4.0), which allows third parties to use the publication as long as the work's authorship and its first publication in this journal are mentioned.