DOCUMENTOS DE TRABAJO

BILTOKI

D.T. 2010.03

Using discrete choice experiments for environmental valuation.

David Hoyos Ramos

eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Facultad de Ciencias Económicas.
Avda. Lehendakari Aguirre, 83
48015 BILBAO.

Documento de Trabajo BILTOKI DT2010.03
Editado por el Departamento de Economía Aplicada III (Econometría y Estadística) de la Universidad del País Vasco.

Depósito Legal No.: BI-755-2010
ISSN: 1134-8984
URN: RePEc:ehu:biltok:201003

Using discrete choice experiments for environmental valuation

David Hoyos

Environmental Economics Unit, Institute for Public Economics, University of the Basque Country

E48015 Bilbao, Spain

e-mail: david.hoyos@ehu.es

fax: +34946017100

Abstract

This paper provides with a review of the state of the art of environmental valuation with discrete choice experiments (DCE). The growing body of literature on this field serves to emphasise the increasing role that DCE are playing in environmental decision making in the last decade. The paper attempts to cover the full process of undertaking a choice experiment, including survey and experimental design, econometric analysis of choice data and welfare analysis. The research on this field is found to be intense, although many challenges are put forward (e.g. choice task complexity and cognitive effort, experimental design, endogeneity or model uncertainty). Reviewing the state of the art of DCE serves to draw attention to the main challenges that this methodological approach will need to overcome in the coming years and to identify the frontiers in discrete choice analysis.

Keywords: discrete choice experiments, choice modelling; survey; environmental valuation

JEL: Q51

Using discrete choice experiments for environmental valuation

1. Introduction

Initially developed by Louviere and Hensher (1982) and Louviere and Woodworth (1983), DCE resulted from the advances in many different disciplines: axiomatic conjoint measurement and information integration theory in psychology, random utility theory-based discrete choice models in economics, and discrete multivariate models for contingency tables and optimal experimental design in statistics (Lancsar and Louviere, 2008). The first application of a DCE in the context of environmental resources was reported by Adamowicz et al. (1994). In the last decade the number of applications has significantly increased and DCE have become a popular stated preference method for environmental valuation.

Discrete Choice Experiments (DCE) or, more generally, Attribute Based Stated Choice Methods (ABSCM) involves the generation and analysis of choice data through the construction of a hypothetical market using a survey. DCE consist of several choice sets, each containing hypothetical alternatives between which respondents are asked to choose their preferred one. Alternatives are defined by a set of attributes, each attribute taking one or more levels. Levels describe ranges over which attributes vary across alternatives. Individuals' choices imply implicit trade-offs between the levels of the attributes in the different alternatives included in a choice set. A baseline alternative (sometimes referred as status quo or 'do nothing' option) is usually included because one of the alternatives must always be in the respondents' feasible choice set so that the results can be interpreted in standard welfare economic terms (Hanley et al., 2001).[1] When the cost or price of the programme is included as an attribute, marginal utility estimates can easily be converted into WTP estimates for changes in the attribute levels and, by combining different attribute changes, welfare measures may be obtained. Furthermore, given that compensating variation measures are obtained, results can directly be used within the cost-benefit analysis framework. Experimental designs are used to construct the choice sets, so that the attributes are uncorrelated and therefore yield un-confounded estimates of the parameters. The resulting choices are finally analysed to estimate the contribution that each attribute and level add to the overall utility of individuals.

The economic model underlying a DCE is intrinsically linked to the statistical model: it conditions the design of the survey and the analysis of data. As a consequence, undertaking a DCE can be gathered as an integrated and cyclical process in which an economic model describing the issue under analysis is permanently revised as new information is gathered from the experimental design, experts' advice, focus groups and pilot surveys (Alpizar et al., 2001).

Different elicitation methods involve a different choice task for individuals to perform (Adamowicz et al., 1998). Hence, ABSCM can take different forms: choice experiments, contingent ranking, contingent rating, paired comparisons and contingent grouping (see Table 1). In the DCE, respondents are presented with a base state situation and some alternative states of the resource under valuation and are asked to choose the option that maximises their welfare. The contingent ranking method is similar but, instead of choosing, respondents are asked to rank the alternatives from the least to the most preferred one. This method provides with more information about the preferences of the individuals but it also adds cognitive difficulty to the choice task. To conform to standard consumer theory, rankings chosen after

[1] If the status quo option is not included, respondents would be 'forced' to choose. As a consequence, the estimates of consumer welfare would be inaccurate and, more importantly, inconsistent with demand theory.

the status quo should be discarded from the estimation procedure.[2] In the contingent rating method individuals are asked to rate in a given scale the different alternatives presented. In this case, the researcher obtains a measure of the magnitude and intensity of the utility associated with each alternative presented. Applications of this methodology in environmental economics are scarce due to the difficulty that embraces transforming ratings into utilities and, more importantly, because of its inconsistency with consumer theory.

Finally, the paired comparison is a variant of the DCE in which alternatives are presented in pairs to the respondent, who is not only asked to choose her preferred option but to indicate the strength of her preferences. More recently, a contingent grouping approach has also been proposed (Brey et al., 2007). In this method, individuals are asked for sorting the alternatives as better than or worse than the base state situation.

Table 1 | Some ABSCM alternatives

Approach	**Task**	**Welfare consistent estimates?**
Choice experiments	Choose between two or more alternatives (where one is the status quo)	Yes
Contingent ranking	Rank a series of alternatives	Depends
Contingent rating	Score alternative scenarios on a scale of 1-10	Doubtful
Paired comparisons	Score pairs of scenarios on a similar scale	Doubtful
Contingent grouping	Sort alternatives as better or worse than the status quo	Yes

Source: Adapted from Hanley et al. (2001)

At the cost of a higher cognitive burden, DCE have demonstrated to have some advantages over the contingent valuation method (CVM): (1) they may reduce potential biases of CVM; (2)

[2] Caparros et al. (2008) show that DCEs and ranking experiments may provide similar results when the first rank is taken as a preferred option.

they are capable of capturing more information from each respondent; and (3) they can test for internal consistency (Alpizar et al., 2001). A comprehensive overview of this valuation method can be found in Ben-Akiva and Lerman (1985), Louviere, Hensher and Swait (2000), Train (2003) and Hensher, Rose and Greene (2005).

This paper aims to review the state of the art of environmental valuation using DCE. Two further objectives are: to draw attention to the main challenges that this methodological approach will need to overcome in the coming years and to identify the frontiers in discrete choice analysis. The paper will be structured as follows: Section 2 presents the underlying economic theory of DCE; Section 3 describes the process of designing a DCE; Section 4 presents the econometrics of discrete choice responses; Section 5 discusses issues related to discrete choice analysis and interpretation; Section 6 outlines the frontiers in discrete choice analysis and Section 7 provides some concluding remarks.

2. Economic theory

The foundation for most microeconomic models of consumer behaviour is utility maximisation under a budget constraint. The theoretical basis of DCE can be found in Lancaster's (1966) characteristics theory of demand, welfare theory and consumer theory, in which individuals derive utility from the characteristics of a good rather than from the whole good. The DCE approach to preference elicitation is similar to the choice-based approach to consumer theory because it is explicitly assumed that respondents' observed choices in the experiment reveal the preferences of the individuals. Thus, DCE combine the Lancasterian theory of value and the consumer demand models developed by Hanemann (1984). According to Hanemann (1984), the consumer decision can be separated into a discrete/continuous choice: which good to choose and how much of the chosen good to consume. In the context of a choice experiment, the decision is constructed so that the discrete choice is isolated. Given the public nature of many non-market goods (mainly in the sense that the quantity is fixed for all agents) each individual is able to choose only one alternative of the choice set, considering both its cost and its continuous dimension.

More formally, each individual is assumed to solve the following maximisation problem:

$$Max_{c,\delta,x} U[\delta_1 c_1(A_1), \delta_2 c_2(A_2), ..., \delta_N c_N(A_N); z] \tag{2.1}$$

s.t. (i) $y = \sum_{i=1}^{N} p_i \delta_i c_i(A_i) + z$

(ii) $\delta_i \delta_j = 0, \forall i \neq j$

(iii) $z \geq 0, \delta_i \geq 0$ for at least one *i*,

where, *U[…]* denotes a quasiconcave utility function; δ_i is a dichotomous variable equal to one if the alternative *i* is chosen and zero otherwise; $c_i(A_i)$ is the alternative combination *i* as a function of its attributes, the vector A_i; p_i is the cost/price attribute of each alternative; *y* is the level of income; and *z* is a composite bundle of goods with its price normalised to 1.

The maximisation problem specified has some properties (Alpizar et al., 2001):

1) All the relevant alternatives are defined and described by all the relevant attributes. Therefore, the selection of attributes and attribute levels has a direct impact on the utility function defined.

2) The price variable in the budget constraint must be related to the full set of attributes conforming each alternative, thus reflecting a continuous dimension.

3) Restriction *(ii)* implies that only one alternative can be chosen in a given choice set.

4) For a given income level, the selection of one alternative (provided in an exogenously fixed quantity) implies that the amount of ordinary goods that can be purchased is also fixed.

5) Restriction *(iii)* implies that the individual will choose a non-negative quantity of the composite good and that an opt-out or status quo option is given.

Two further assumptions need to be made in order to solve the maximisation problem: firstly, a purely discrete choice is assumed, and secondly, weak complementarity is assumed (i.e. the attribute levels of the non-selected alternatives have no influence on the utility function of the chosen alternative). Following Hanemann (1984):

$$\text{If } \delta_i = 0 \text{, then } \frac{\partial U}{\partial A_i} = 0 \ \forall i \neq j \ . \tag{2.2}$$

Following Equations (2.1) and (2.2), and given $\delta_j = 1$, the conditional utility function can be written as:

$$U_j = V_j \lfloor c_j(A_j), p_j, y, z \rfloor = V_j \left(A_j, y - p_j c_j \right). \tag{2.3}$$

Going back to the unconditional indirect utility function, given by the following expression:

$$V[A, p, y] = \max[V_1(A_1, y - p_1 c_1), \ldots, V_1(A_N, y - p_N c_N)], \tag{2.4}$$

it follows that, for a purely discrete choice, as the one described above, the individual chooses alternative *j* if and only if:

$$V_j(A_j, y - p_j c_j) > V_i(A_i, y - p_i c_i), \forall i \neq j \ . \tag{2.5}$$

It is important to denote that the model above is a generalisation of the economic model underlying a close-ended CV survey. In the CVM there would only be two alternatives: the situation before the project and after the project so that an individual will accept the bid if her utility increases. Equations (2.4) and (2.5) also provide the deterministic model of consumer behaviour that will serve as a basis for the econometric model and the estimation of welfare effects that will be discussed in the following Section. In order to be operational, two decisions will need to be taken: the functional form of the utility function and the distribution of the error term that will be introduced in the model in order to capture unobservable behaviour.

3. Designing a DCE

3.1. Questionnaire development

Designing and implementing a DCE requires a proper survey design. All the recommendations available for CV surveys (e.g. Mitchell and Carson (1989) among others) are also applicable to DCE. However, special attention needs to be put in the conceptualisation of the choice process. Two issues may arise at this point: first, the analyst should pursue an incentive compatible choice question to avoid respondents to not giving their true preferences; secondly, the choice format should mimic as much as possible the actual choice context. Another focus of attention is the existence of a status quo option, especially in order to derive proper welfare measures. Designing a DCE is a cyclical process involving four steps: (1) definition of attributes and levels of provision, (2) experimental design, (3) questionnaire development, and (4) sampling strategy. Feedbacks from different stages are sequentially incorporated in the final design of the DCE.

Environmental attributes and level of provision become critical aspects of any DCE given that the only information about preferences provided by respondents takes the form of choices between these options (Hensher, 2007). Attributes can be quantitative or qualitative and can be generic (same levels for all alternatives) or alternative specific (some attributes or levels may differ across alternatives). According to Lancaster (1991), an environmental attribute can be considered relevant if ignoring it would change our conclusions about the preferences of consumers. While respondents may consider relevant a different set of attributes, it is important that the DCE captures the main attributes to the majority of respondents so concerns about omitted attributes are avoided. The construction of the choice sets included in an experiment requires a correct definition of the change to be valued and the attributes and levels that would be used. Previous scientific investigation on the environmental characteristics of the good or service under valuation, expert advice and focus groups may facilitate the definition of attributes and levels of provision. Focus groups may also help in deciding the best strategy for explaining the task of making successive choices from a series of choice sets.

Levels should be plausible and relevant but this does not mean that they need to be currently available. Focus group sessions and pilot surveys may also help to identifying the appropriate levels of cost attribute. The payment vehicle and duration should be chosen in consonance with the good under valuation and its context. As Louviere et al. (2000) argue, the suitable number of attributes and levels is context specific. The analyst should weight up the relevant number of attributes and the complexity of the design. The trade-off between the possibility of omitted variable bias and task complexity and cognitive burden to respondents may be analysed in focus groups and pilot surveys. Additionally, it may be interesting to use focus groups to identify any possible interaction effect between attributes. Complexity of the choice task can be investigating with verbal protocols (Schkade and Payne, 1994). It may be useful to provide “warm-up” choice tasks to ensure repospodents’ correct understanding of the task (Carson et al., 1994).

Finally, iterative pilot tests are required in order to develop a DCE survey. Pilot tests should check for respondents’ understanding of the choice context and task, the adequacy of the attributes and levels considered and other factors such as length and timing.

3.2. Experimental design

The DCE data generation process relies on experimental design. An experimental design is a combination of attributes and levels used to construct the alternatives included in the choice sets. The generation of the experimental design represents a main and complex component of stated choice studies. In fact, the allocation of attributes and attribute levels in generating choice sets for an experimental design can have a significant influence on the study outputs (Bliemer et al., 2009). In a typical DCE, respondents face a number of hypothetical scenarios (choice sets) containing a set of alternatives that differ on a number of attribute or level dimensions. Respondents' stated alternative choices in every choice situation are used to estimate parameter weights for each of the attributes and the analyst may obtain estimates of marginal rates of substitution between them. Identification and efficiency are the two main statistical issues involved in the experimental design construction. Identification is related to the effects that can be independently estimated, which is further related to the specification of the indirect utility function. Efficiency, on the other hand, is related to the precision of the parameter estimates.

So, before creating an experimental design, the model and the parameters to be estimated need to be specified. Essentially, this means that the first step involves a complete specification of the utility functions. For this purpose, it is important to consider the number of alternatives and the number of attributes that will form each alternative, the consideration of generic or alternative-specific attributes (including alternative specific constant issues), the inclusion of interaction effects between attributes and the consideration of nonlinear effects via dummy-coded or effects-coded variables. At this stage, it is important to bear in mind two issues (Bliemer and Rose, 2006): firstly, that each additional parameter represents an extra degree of freedom. In other words, the number of choice sets in the experimental design must be equal or greater than the degrees of freedom (i.e. the total number of parameters excluding the constants plus one); and, secondly, that the design would be suboptimal if at the estimation stage the model estimation is changed or extra variables are added to the utility functions (e.g. socioeconomic variables).

The form of the estimated indirect utility function will ultimately depend on the experimental design and the type of choice model used. Different choice models can be estimated depending on the assumption about the distribution and properties of the error component and about the variance-covariance matrix of estimated preference parameters. Sometimes economic theory may give some advice on the functional forms of individual variables. Alternatively, some authors recommend estimating the most disaggregated possible model by including parameters estimates for every attribute level but one, and mapping obtained parameters against the attribute level to visualise its functional form.

The second main step in the experimental design construction involves the generation of the experimental design. Prior to the design some decisions need to be taken: (1) whether the experiment will be labelled or unlabelled; (2) the consideration of attribute level balance[3] so that it is ensured that the parameters can be estimated on the whole range of levels; (3) the number of attribute levels, considering that the more levels used and the more different the levels are between the attributes the higher number of choice sets will be; and (4) the attribute level range should be wide enough so that parameter estimates will have smaller standard errors and it ensures a broader application interval.[4]

At this stage, several different designs can be considered. A full factorial design includes all possible combinations of attributes and levels. Given that all possible combinations are included, a full factorial design allows for estimation of main effects and interactions effects independently of one another. However, given that the number of combinations may become too large, fractional factorial designs are usually implemented. A fractional factorial design is a sample of the full design that allows the estimation of all the effects of interest (typically main

[3] Attribute level balance requires that the different levels appear an equal number of times for each attribute.

[4] It is important to bear in mind that the model estimated will only be applicable on the data range it was estimated on.

effects only or main effects plus some higher-order interaction effects). Full and fractional factorial designs can also be blocked into different versions to which respondents are randomly assigned. In case of a blocked design some authors have suggested to include a version variable in the estimation to account for version effects (Lancsar and Louviere, 2008). Fractional factorial designs can be orthogonal (i.e. those pursuing no correlation between the attribute levels) or so-called efficient designs (i.e. those pursuing the minimum predicted standard errors of the parameter estimates). An additional difference between the two is the amount of information required because efficient designs rely on prior information about the parameter estimates.

Efficiency is a measure of the level of precision in which effects are estimated. Various efficiency criterions have been proposed, such as A-error or D-error. The D-error has become the most widely used measured of efficiency because of its insensitivity to the magnitude of the scale of the parameters. According to Street et al. (2005), D-efficiency is determined by the following expression:

$$D - Efficiency = \left[\det(C) / \det(C_{opt})\right]^{1/\rho}, \qquad (3.1)$$

where C is the Fisher Information Matrix and C_{opt} is the largest value of C, *det* stands for determinant and ρ is the number of parameters to be estimated in the model.

Huber and Zwerina (1996) distinguish four characteristics for an efficient experimental design: (1) orthogonality; (2) level balance; (3) minimal overlap; and (4) utility balance. As mentioned above, the problem with so-called efficient designs is that they require prior information about the true distribution of the parameters. Optimal efficient design is a research field in constant progress during the last few years.

Although the inclusion of a status quo option may reduce efficiency, it may be justified on the grounds of better congruency with consumer theory and real choices. Furthermore, in non-labelled experiments the inclusion of status quo options does not affect optimality. Deshazo and Fermo (2002) argue that as the complexity of DCE increases choice inconsistency increases, although careful design and estimation may help mitigating unobserved variability. The empirical evidence on task complexity suggests that experiments should be very carefully designed and estimated, and that it should be no more complex than the market it aims to simulate. If the analyst needs to handle implausible attribute combinations in order to increase realism (either by constraining the design or by randomly substituting implausible for plausible combinations) she should acknowledge the subsequent loss of efficiency and make sure that the properties of the design remain desirable. In other words, there may be a trade-off between optimality and plausibility. It is clear that from a statistical perspective optimal design is desirable, but from an empirical perspective some other issues need to be taken into account, such as task complexity, heuristics or the inclusion of a base scenario or status quo option (Lancsar and Louviere, 2008).[5]

One important question that arises when designing the price vector of an experimental design is the sensitivity of the WTP estimates to changes in the design. In other words, the question is whether DCE would share the anchoring, starting-point of framing effects sometimes encountered in CVM. The empirical evidence suggests, however, that when accounting for scale effects the structure of preferences and WTP estimates is not significantly different when changing the price vector (Hanley et al., 2005). These issues will be further analysed in Section 5.4.

Computer software, including SAS, SPSS, JMP or more recently NGENE, has been used to generate experimental designs. Other resources on the internet include a library of orthogonal

[5] In the presence of complex choices respondents may use heuristics or rules of thumb to simplify the decision task, including maximin and maximax strategies and lexicographic ordering (Tversky, 1972).

arrays[6] as well as a web page devoted to constructing or checking designs developed by Devorah Street and Leonie Burgess.[7]

3.3. Survey administration

The proper sampling strategy requires a consideration of the relevant population. Sample size should be as large as the requirements of the estimation of reliable models, but obviously subject to available budget and other constraints. It is difficult to determine the optimal sample size for non-linear choice models because it depends on the true values of the unknown parameters estimated.

Several sampling strategies can be adopted. In a simple random sampling, the probability of being drawn is identical for each individual. In an exogenously stratified sampling, the probability of being drawn depends on independent variables. Finally, in an endogenously stratified sampling the probability of being drawn depends both on dependent and independent variables. Estimation methods vary over the different sampling procedures. If the sample is exogenously drawn (either random or stratified random), maximum likelihood estimation may be used. If, on the other hand, the sample is endogenously drawn, conditional maximum likelihood estimation is required (Train, 2003).

4. Econometrics of discrete choice experiments

It is not infrequent to find that observed choices may reveal preference structure inconsistent with the deterministic model described in the previous Section. It is generally assumed that these inconsistencies are due to unobserved factors such as characteristics of the individual,

[6] http://www.research.att.com/~njas/oadir/index.html

[7] http://crsu.science.uts.edu.au/choice/choice.html

non-included attributes in the experiment, measurement error or heterogeneity of preferences (Hanemann and Kanninen, 1999).

4.1. Models for discrete choice analysis

The analysis of the choices made in DCE is based on random utility theory. The random utility approach developed by McFadden (1974) is used to link the deterministic model with a statistical model of human behaviour. The randomness of the utility function suggests that only analysis of the probability of choosing one alternative over another is possible. Estimable choice models require a distributional assumption for the random component.

As a consequence, a random disturbance with a specified probability distribution, ε, is introduced into the economic model. In this context, an individual will choose alternative *j* if and only if:

$$V_j(A_j, y - p_j c_j, \varepsilon_j) > V_i(A_i, y - p_i c_i, \varepsilon_i), \forall i \neq j \,. \tag{4.1}$$

Or, in probability terms (following the notation from the previous Section):

$$P(\delta_j = 1) = P\big(V_j(A_j, y - p_j c_j, \varepsilon) > V_i(A_i, y - p_i c_i, \varepsilon)\big), \forall i \neq j \,. \tag{4.2}$$

The final specification of the econometric model will ultimately depend on two further decisions: (i) the specification of the utility function (i.e. how the random term enters the conditional indirect utility function), and (ii) the distributional assumption for the error component.

The most common formulation for the utility function is additively separable, so that the error component enters the utility function as an additive term. Under this assumption, the probability statement of Equation (4.2) becomes:

$$P(\delta_j = 1) = P\big(V_j(A_j, y - p_j c_j) + \varepsilon > V_i(A_i, y - p_i c_i) + \varepsilon\big), \forall i \neq j\,. \qquad (4.3)$$

The specification of the utility function requires two additional decisions: the functional form for V_j and the relevant attributes A_j that will determine the utility level for each alternative. The deterministic component of utility is usually assumed to be a linear and additive function of the attributes of the good and the characteristics of the respondent (X_j):

$$V_j = \beta' x_j\,, \qquad (4.4)$$

where β is a vector of coefficients to be estimated. Utility is a latent quantity while choices are the only observable indicator of utility. The linear in parameters formulation of the utility function is simple and convenient, and it does not exclude the possibility of including non-linear effects on utility through for example a quadratic utility function. Less restrictive formulations (although sometimes more flexible) may carry some complications. One crucial assumption relates to the way that income enters the utility function. A constant marginal utility of income is usually assumed because it facilitates the estimation of welfare measures, although it may not be always reasonable. Regarding the influence of selected attributes and interactions it is important to denote that the data collected in a DCE is based on a specific experimental design that will condition the estimation of interaction effects between the relevant attributes.

The second decision regarding the final specification of the econometric model relates to the specification of the probability distribution of the error term. Assuming that the DCE has M choice sets (S_m), each formed by K_m alternatives with A_i attributes such that $S_m = \{A_{1m}, \ldots, A_{Km}\}$, the probability of choosing alternative j from a choice set S_m can be written as:

$$P(\delta_j = 1 | S_m) = P\begin{Bmatrix} V_j(A_{jm}, y - p_j c_j) + \varepsilon_j > \\ V_i(A_{im}, y - p_i c_i) + \varepsilon_i; \forall i \in S_m; \forall i \neq j \end{Bmatrix}. \quad (4.5)$$

Under the assumption that the error terms of the utility function are independently and identically distributed following a type I extreme value (Gumbel) distribution, the choice model can be estimated using a multinomial logit (MNL) specification (McFadden 1974, Louviere et al. 2000). In general, utility can be expressed as $U_{ij}^* = V_{ij} + \varepsilon_{ij}^*$, where the error component has variance $\sigma^2 \times (\pi^2/6)$. Thus, utility may be divided by σ without changing behaviour given that the scale of utility is irrelevant to behaviour: $U_{ij} = V_{ij}/\sigma + \varepsilon_{ij}$, where $\varepsilon_{ij} = \varepsilon_{ij}^*/\sigma$ and the variance of the error component simply becomes $\pi^2/6$. The MNL statistical model represents the probability of choosing an alternative *j* such that the utility of that alternative is greater than the utility of all other alternatives. The probability of an individual *i* choosing an alternative *j* is:

$$P_{ij} = \frac{e^{V_{ij}/\sigma}}{\sum_{h \in c} e^{V_{ij}/\sigma}}. \quad (4.6)$$

Given that we have assumed that V_{ij} is linear in parameters with coefficients β^*, the choice probability can be rewritten as:

$$P_{ij} = \frac{e^{(\beta^*/\sigma)' x_{ij}}}{\sum_{h \in c} e^{(\beta^*/\sigma)' x_{ij}}}, \quad (4.7)$$

so that each coefficient is in fact scaled by $1/\sigma$ (and therefore σ is called the scale parameter) and and *c* is a choice set. In other words, the scale parameter scales the *true* parameters to reflect the variance of the unobserved portion of utility, so that the parameter estimates β are, in fact, β^*/σ.

The scale parameter cannot be identified from the data set so it is usually normalized to one, implying constant error variance. However, the existence of a scale parameter has two implications for the analysis of results: on the one hand, an increase in the scale reduces the variance (i.e. high fitting models have larger scales); and secondly, it imposes restrictions on the interpretation of the estimated coefficients (i.e. coefficients can be compared within an estimated model but not across different models).

The key assumption of the MNL model is not so much that the error term is distributed following a Gumbel distribution, but that the errors are independent of each other. It basically means that the unobserved part of utility for one alternative is unrelated to the unobserved part of utility for the other alternatives. If this is not the case, the researcher faces three options (1) to specify a model that allows for correlation, (2) to respecify the utility function so that the source of correlation is captured by another variable and the error terms remain independent, or (3) to maintain the current specification of utility acknowledging that the model may be an approximation.

The power and limitations of the MNL model can be elucidated under the following criteria (Train, 2003): (1) MNL can represent systematic taste variations (i.e. those related to observed characteristics of the respondents) but not random taste variations (i.e. those that cannot be linked to observed characteristics of the respondents), (2) the MNL exhibits restrictive substitution patterns because it implies proportional substitution across alternatives given the specification of the utility function, and (3) the MNL can handle situations where unobserved factors are independent over time but it cannot be used with panel data when unobserved factors are correlated over time for each respondent.

In the first place, observed heterogeneity can be incorporated into the systematic part of the model by including interactions between attributes or constant terms and socioeconomic characteristics of the respondents. Therefore, the MNL model can only handle observed heterogeneity but not unobserved heterogeneity.

The substitution patterns of MNL models relate to the property of Independence of Irrelevant Alternatives (IIA). The IIA axiom states that “the ratio of the probabilities of choosing one alternative over another (given that both alternatives have a non-zero probability of choice) is unaffected by the presence or absence of any additional alternatives in the choice set” (Louviere et al., 2000):

$$\frac{P_{ij}}{P_{ik}} = \frac{e^{V_{ij}}}{e^{V_{ik}}}\,. \tag{4.8}$$

As a consequence, IIA depends both on the choice and on the variables included in the specification of V_{ij}, that are assumed to be identically and independently distributed (IID). In case of violation of IIA, the parameters estimation would be biased. The IIA property is usually checked using the test proposed by Hausman and McFadden (1984). Alternatively, IIA can be more simply tested by estimating a mixed logit model and testing whether the variance of the mixing distribution is equal to zero.

In the context of panel data, where several choices from the same individual are observed, the MNL specification treats the data as cross-sectional given that the unobserved factors affecting the respondent’s choice are assumed to be independent.

The independence of the error terms assumption may seem restrictive but, in fact, it can be seen as the natural outcome of a correctly specified utility function that captures all sources of correlation explicitely and leaves as a ‘white noise’ the unobserved part of utility (Train, 2003). However, not infrequently the IIA hypothesis does not hold and the researcher needs to incorporate correlation among alternatives. Generalised extreme value (GEV) models are a bunch of models attempting to include different substitution patterns under a unifying framework: unobserved utility for all alternatives are jointly distributed as generalised extreme value. These models, therefore, relax the second restriction of the MNL model. GEV models collapse to the standard MNL model when all correlations are zero. The most widely used GEV model is the nested logit model, although other suggested models include the cross-nested logit model, the network GEV model, the paired combinatorial logit model or the generalised nested logit model.

Multinomial Probit (MNP) models deal with all three previously detected restrictions: they can incorporated random taste variations, they allow any pattern of substitution and they can handle with correlated errors. However, this model specification has two limitations: firstly, it is computationally more demanding and, more importantly, it requires normal distributions for all unobserved components of the utility function. The problem is that, under some circumstances, normal distributions may not be appropriate.

The Mixed Logit (MXL) model is able to overcome all the limitations previously enumerated. It is important to denote that the MXL model nests many particular specifications used in the relevant literature. Furthermore, under some basic conditions, the choice probabilities of any RUM discrete choice model can be derived from a MXL model specification (McFadden and Train, 2000).

The remarkable growth in the use of MXL models in recent years can be partly explained by their inclusion in standard econometric software and partly by their flexible assumptions. There are three main advantages when a MXL model specification is used: preference heterogeneity is directly incorporated through individuals' random taste variations; it avoids any reliance on the IIA property; and it is capable of incorporating correlation across choice sets and alternatives. Its popularity has kept growing in spite of some problems related to inference and model selection (Brownstone, 2001).

MXL is a mixture of the logit function (see Equation (4.6)) evaluated at different parameters (β) with $f(\beta)$ as the mixing distribution:

$$P_{ij} = \int \left(\frac{e^{\beta' x_{ij}}}{\sum_{h \in c} e^{\beta' x_{ij}}} \right) f(\beta) d\beta \, . \tag{4.9}$$

The mixing distribution of the parameters can be discrete or continuous. If the mixing distribution is formed by a finite set of distinct values, the MXL becomes the latent class model. If the mixing distribution is continuous, a random parameters logit (RPL) model (also known as random coefficients model) or an error component (EC) model can be derived from the MXL probability.

In the RPL model, a random term whose distribution over individuals and alternatives depends in general on underlying parameters is added to a classical utility function associated with each alternative, that is:

$$U_{ij} = \beta' x_{ij} + \eta_i x_{ij} + \varepsilon_{ij} \, , \tag{4.10}$$

where η_i is a vector of deviation parameters and ε_{ij} is the error component. If the error terms are IID type I extreme value, we have a random parameter logit (RPL) model; if the error terms are IID normal distributed, we have a random parameter probit (RPP) model. So, the RPL is a model in which an individual's utility from any alternative in the choice set includes a stochastic part that may be correlated over alternatives and that may be heteroskedastic (Henser et al., 2005). In this model, preference heterogeneity is directly incorporated into the vector of parameters, so that the vector of coefficients of attributes is different for each individual, β_i, and it is allowed to deviate from the population mean coefficient β by the vector of deviation parameters η_i.

The use of a MXL model involves three main specification issues: (1) the determination of which parameters should be modelled as randomly distributed, (2) the choice of mixing distribution for the random coefficients and (3) the economic interpretation of the randomly distributed coefficients.

The classical procedure to determine the random coefficients is to select among different model specifications (including/excluding random coefficients) using the Likelihhod Ratio (LR) test. A second possibility is the use of the Lagrange Multiplier (LM) test, as proposed by McFadden and Train (2000). This test proceeds by constructing artificial variables:

$$z_{tj} = \frac{1}{2}\left(x_{tj} - x_{tC}\right), \tag{4.11}$$

where $x_{tC} = \sum_{j \in C} x_{tj} P_j$, *t* denotes the parameters that are suspected to be random, *C* is the set of alternatives being offered, and P_j is the choice probability for alternative *j*. Once the model is reestimated including the artificial variables, the null hypothesis of fix parameters for *t* can be rejected using a simple LR or Wald test when the coefficients for the artificial variables are significantly different from zero. The choice of the mixing distribution is part of an ongoing debate that will be further analysed in Section 6.3.

Alternatively, a MXL model can be derived without a RPL interpretation, as simply representing error components creating correlation among the utilities for different alternatives. In this case, the unobserved (random) part of utility is becomes: $\eta_{ij} = \mu_i x_{ij} + \varepsilon_{ij}$, so that correlation over alternatives depends on the specification of x_{ij}. If there is no correlation among alternatives, x_{ij} is identically zero and the probability model collapses to the standard MNL model.

Finally, it is also important to bear in mind that in the MXL it is further assumed that the preferences are allowed to vary among individuals but are stable among the choice situations, although this assumption can be relaxed to allow for fatigue or learning effects.

4.2. Welfare measures

Following standard consumer theory, the marginal rate of substitution (MRS) between attributes can be obtained by calculating the ratio of the partial derivatives of the indirect utility function with respect to each attribute:

$$MRS_{X_1,X_2} = \frac{\partial V / \partial X_1}{\partial V / \partial X_2}. \tag{4.12}$$

In the presence of a linearly additive indirect utility function as in Equation (4.4), the MRS equals the ratio of the estimated parameters. If the second attribute is the cost attribute, it is usually termed 'implicit price'.[8] Similarly, compensating surplus (CS) welfare estimates for DCE may be obtained from Hanemann (1984b) and Train (1988):

$$CS = -\frac{1}{\alpha}\left[\ln(\sum \exp(\beta X_{ij}^{0}) - \ln(\sum \exp(\beta X_{ij}^{1})\right], \tag{4.13}$$

where α is the marginal utility of income (usually represented by the coefficient of the payment attribute) and X_{ij}^{0} and X_{ij}^{1} represent the vector of environmental attributes at initial level (status quo) and after the change levels, respectively. So Hicksian compensating variation measures a change in expected utility due to a change in the level of provision in the attribute or attributes by weighting this change by the marginal utility of income. Simplifying the above Equation, the marginal value of a change in one attribute is measured through the ratio of the two coefficients:

[8] It is important to bear in mind that marginal rates of substitution measures such as WTP are not affected by the scale parameters because these cancel out.

$$WTP = \frac{\beta_{attribute}}{\beta_{\cos t}}. \tag{4.14}$$

Therefore, the WTP for a marginal change in the level of provision of each environmental attribute (i.e. the marginal rate of substitution between income change and this attribute change) is obtained by dividing the coefficient of the attribute by the coefficient of the cost attribute. It is not infrequent to report not only the marginal WTP but the distribution of welfare effects. This is usually done by bootstrapping techniques such as the Krinsky and Robb (1986) procedure.

Finally, it is important to bear in mind the assumptions underlying the closed-form solution for the welfare measure in Equation (4.13) being: additive disturbances, an extreme-value distribution and constant marginal utility of income. The problem of relaxing the hypothesis of constant marginal utility of income is that it complicates the estimation of compensating surplus measures because income enters the utility function non-linearly. Some approaches to incorporate income effects in random utility models have been proposed by McFadden (1995) and Morey et al. (1993).

5. Discrete choice analysis

The development of a model specification is not straightforward. Instead, it can be gathered as a combination of behavioural theories and statistical methods with subjective judgements of the researcher. The ‘art of model building’, as some authors argue, suggests that the most appropriate specification of a DC model requires formal theories, researcher’s *a priori* knowledge of the phenomenon being modelled and statistical inference. In practice, the model development process generally begins with some theoretical inputs. It follows with a sequential process of model estimations and various formal and informal tests for narrowing the alternative specifications. In this process, the analyst tries to revise the *a priori* assumptions with the statistical findings. Once the model specification is found to be consistent with the theory, model selection is based on goodness-of-fit performance and statistical significance tests.

5.1. Model specification

In the process of model specification, the analyst needs to tackle many important issues. In the previous Section, models for DCE have been reviewed. The present Section will focus on two related issues: the coding of explanatory variables and the inclusion or exclusion of alternative specific constants (ASC).

There are three options for the coding of explanatory variables: mean-centering numerical attributes, effects coding and dummy coding. The first two options are usually preferable because they avoid correlation with the intercepts and minimise collinearity in estimation matrices used to estimate interactions. Even though effects coding estimation requires omitting one level, it can be easily calculated as minus one times the sum of the estimated levels.

The inclusion or exclusion of ASC in DCE has also received special attention. ASC refers to a parameter for a particular alternative representing the role of unobserved sources of utility. The problem is that the inclusion or omission of elements of the utility function in the estimation process can have a significant influence on the obtained welfare measures (Mogas et al., 2006). In the context of unlabelled experiments (i.e. those experiments containing generic alternatives), such as those generally used in environmental valuation, it has been argued that including an ASC would *violate* the meaning of unlabelled and that the correct way to proceed would be to exclude constant terms for all unlabelled alternatives (Hensher et al., 2005). By doing this, the average unobserved effect for all alternatives is constrained to be zero. The omission of the ASCs has been justified on the grounds that, even thought they may improve the model fit, they are not related to specific attributes and hence do not explain choices in terms of observable attributes (Adamowicz et al., 1997). In other words, they lack behavioural interpretation. Applications excluding ASCs are abundant in the literature (Bienabe and Hearne, 2006; Nielsen et al., 2007; Xu et al., 2003). However, when excluding an ASC the remainder of the model parameters would attempt to capture this effect, resulting in biased attribute parameter estimates. Hence, it has been argued that ASCs are important in order to interpret the preferences of the individuals (Morrison et al., 2002). For instance, it has been interpreted as a status quo bias[9] or endowment effect (Adamowicz et al., 1998). Alternatively, depending on the sign, it has also been interpreted as that respondents may have a utility premium for moving away from the status quo (Mogas et al., 2006). Thus, some studies would use an alternative specific constant for the status quo alternative, even if the attributes are generic (Birol et al., 2006; Horne et al., 2005; Rolfe et al., 2000). Current state of the art in discrete choice analysis favours the latter approach.

5.2. Goodness of fit

Goodness of fit of estimated models is usually measured by a statistic called the likelihood ratio index (LRI) or pseudo-Rho squared. This statistic is often used to measure how well the estimated model fits the data in the context of DCE. The LRI is defined as:

$$\rho^2 = 1 - \frac{LL(\hat{\beta})}{LL(0)}, \tag{5.1}$$

where LL($\hat{\beta}$) is the value of the log-likelihood function at the estimated parameters and LL(0) denotes its value when all the parameters are set equal to zero. If the explanatory power of

[9] The notion of 'status quo bias' (Samuelson and Zeckhauser, 1988) refers to a common economic phenomenon where survey respodents tend to attach to the current situation.

the model is very low, the estimated model is no better than no model and so the LRI takes the value of zero. On the opposite, if the estimated model perfectly predicted every choice, the LRI would take the value one.

Another informal goodness-of-fit test is the adjusted LRI (rho bar squared). It is similar to the previous statistic but it corrects for the number of parameters estimated:

$$\bar{\rho}^2 = 1 - \frac{LL(\hat{\beta}) - K}{LL(0)}, \tag{5.2}$$

where *K* denotes the number of unknown parameters in the model.

5.3. Model selection

Selection of the appropriate model may be based on economic and behaviour theory as well as on statistical considerations such as LR test in the presence of nested models. The LR statistic is:

$$-2\left[LL(\hat{\beta}_R) - LL(\hat{\beta}_U)\right] \sim \chi^2_{K_U - K_R} \tag{5.3}$$

where $\hat{\beta}_R$ and $\hat{\beta}_U$ denote the estimated coefficients of the restricted and unrestricted models respectively. The statistic is χ^2 distributed with $K_U - K_R$ degrees of freedom, representing the number of estimated coefficients in the unrestricted and restricted models.

While non-nested hypothesis cannot be tested using the LR test, some different options have been proposed. One possibility is to construct a composite model and perform two LR tests for each of the two restricted model against the composite model, known as Cox Test (Cox, 1961; 1962). A second possibility is to compare them using the adjusted LRI statistic and selecting the one with the higher statistic. A third option would be to use the Davidson and MacKinnon J test (Davidson and MacKinnon, 1981).

5.4. Internal and external validity

Internal tests of validity are designed to check the standard assumptions about the preferences of the individuals, namely that preferences are stable, complete, monotonic and transitive. Given that the SP survey designs basically force completeness of preferences, it is generally referred as preference consistency when the monotonicity, transitivity and stability axioms are complied with. The empirical evidence on internal validity is mixed. Carlsson and Martinsson (2001b) find no evidence of violation of the assumptions about transitivity and stability of preferences. On the contrary, many authors have reported preference inconsistency (Deshazo and Fermo, 2002; Johnson and Mathews, 2001; McIntosh and Ryan, 2002). These results seem to confirm the initial worry of many academics that when moving from CVM to DCE, complexity-induced choice inconsistency increases. However, complexity-induced inconsistencies may be mitigated if experiments are designed and estimated more carefully (Deshazo and Fermo, 2002).

Rationality tests play a main role in identifying respondents showing preference inconsistency. The empirical research has shown that including “irrational responses” may bias the point estimates and increase its variance (Johnson and Mathews, 2001; Foster and Mourato, 2003). Rationality tests usually take the form of an additional choice task with a dominant alternative (e.g. Scarpa et al. (2007)). However, the researcher should bear in mind that some factors may underlie what it is labelled as ‘irrational’ and removing all of them may induce sample selection bias and reduce the statistical efficiency and power of the estimated models (Lancsar and Louviere, 2006).

One of the main anomalies detected is failing to comply with the continuity axiom, a basic assumption of DCE whereby there is unlimited substitutability between the attributes employed in an experiment. The implicit assumption in RUM is that individuals' decisions respond to the compensatory heuristics, this is, individuals weight up the relative contribution of the different attributes to overall utility choosing the one reporting the highest utility. However, different disciplines have shown departures from compensatory heuristics such as psychology (e.g. Kahneman and Frederick (2002), Payne et al. (1993) or Scott (2002)) or economics (e.g. Campbell et al. (2008) or Araña and Leon (2009)). The main factors influencing non-compensatory response heuristics are complexity of the choice task and contextual factors.

The discussion above suggests that people may employ simple decision-making heuristics and that it should be considered when deriving welfare estimates in the DCE methodological framework. Process heterogeneity, as defined by Hensher (2008), implies that the discrete choice analysis needs to recognise and account for the many different ways in which individuals process information, in partly influenced by the analyst's description of the context in which preference data is sought. The number of attributes included in a DCE is often related to the complexity of the experiment but, as this author argues, this argument may be misleading since it establishes a direct relationship between complexity and amount of information rather than in relation to the relevance of information.

As a consequence, there is a recent trend aiming to accommodate both the processing of attributes together with the choice outcome under the same framework. One way of dealing with this issue that has received special attention in the recent literature is by modelling attribute non-attendance. Ignoring one or more attributes implies non-compensatory behaviour among respondents (i.e. discontinuous preferences). One example of discontinuous preferences is lexicographic ordering. The problem of discontinuity in the context of DCE is that without continuity marginal rates of substitution between the attributes cannot be computed at the individual level. However, there have been some recent attempts to incorporate discontinuous preferences in the DCE framework by deriving marginal rate of substitution from the estimated parameters at the sampled population level. Pucket and Hensher (2009) and Campbell et al. (2008) find that welfare estimates are biased when discontinuous preferences are ignored. More recently, Meyerhoff and Liebe (2009) find that accounting for discontinuous preferences at the choice task level improves the performance of the estimated models but it does not affect the WTP estimates.

External validity is usually defined as the empirical evidence proving that the choice process and utility estimates obtained in a DCE are similar to those obtained in real markets (i.e. it compares actual and hypothetical behaviour).[10] The close relationship between choice experiments and consumer purchasing behaviour suggests that DCE should pass external validity tests and that they should be less prone to the hypothetical bias sometimes encountered in CV studies. Even though the literature on external validity of DCE is smaller than the one on CVM, it suggests that DCE generally pass external tests of validity (Alpizar et al., 2001; Carlsson and Martinsson, 2001a; Carson et al., 1994). More recently, Lusk and Schroeder (2004) reported the existence of hypothetical bias in average WTP but not in marginal WTP. These results seem to be in line with those obtained in the transportation literature when comparing RP and SP data, where SP data provides reasonable estimates of marginal changes in the quality attributes but it provides poor predictions of actual market shares (Louviere et al., 2000).

Moving from the single binary discrete CV question to the series of questions involved in DCE implies a loss of incentive compatibility that may give rise to strategic behaviour. Although it has been argued that one of the advantages of DCE over CVM is that it may be more difficult to

[10] It should be noted that comparing DCE outcomes requires accounting for differences in scale.

behave strategically in choice experiments,[11] Bateman et al. (2008) find empirical evidence of strategic behaviour in the context of valuing public goods. In respect to the welfare measures obtained, DCE and CVM have been proved to provide similar parameters on the marginal utility of income once variance heterogeneity is taken into account as well as similar error variance (Adamowicz et al., 1998). In general, CVM and DCE seem to provide not significantly different economic values (Adamowicz et al., 1998; Jin et al., 2006a; Mogas et al., 2006).

There has been also some research on the biases sometimes encountered in CV studies. In respect to sensitivity to scope it has been found that DCE show stronger sensitivity to scope than CVM (Goldberg and Roosen, 2007; Foster and Mourato, 2003). Ladenburg and Olsen (2008) find that preferences obtained in DCE are susceptible to the anchoring effects or starting point bias often encountered in dichotomous choice CV studies. These findings suggest that, as a generalisation of the CVM, design issues are as important in DCE as they were in CVM so that a careful survey design can help mitigating the biases sometimes encountered in the literature. However, in order to be well established, DCE need more testing of its properties, as it has been the case with CVM (Bateman et al., 2008).

6. Frontiers in discrete choice analysis

6.1. Specification of the utility function

The most common specification of the utility function in DCE is linear, as it was specified in Equation (4.4). However, it can be argued that utility functions are not likely to be linear due to the existence of diminishing marginal utilities or gain-loss asymmetries. In this context, models capable of accommodating non-linearities in the specification of the utility function are likely to have stronger explanatory variables. For example, an empirical investigation conducted by

[11] For example, Alpizar et al. (2001) argue that by using unlabelled experiments the risk of individuals behaving strategically in DCE may be reduced

Lanz et al. (2009) finds non-linearities both in the environmental and cost attributes that, when non accounted for, end up in overestimation of the WTP and underestimation of the WTA.

Non-linear specifications can take three forms: (1) capturing nonlinearities in the deterministic part of the utility function; (2) inclusion of discrete variables and (3) including multiplicative error terms. Firstly, data can be pre-processed to account for nonlinearities in the utility function (for example, by using the log of a variable instead of the variable itself). Secondly, discrete variables, mainly used to capture qualitative attributes, can be used. And thirdly, two useful approaches for accounting for nonlinearities in the context of continuous variables include the piecewise linear specification and the power series expansion. Nonlinear transformations of variables that are not linear in the unknown parameters include Box-Cox and Box-Tukey transformations (Greene, 2003).

The model specification can be generalised to account for nonlinearities in two ways: by introducing heteroskedasticity in the usual framework or by including random coefficients in the model. Finally, multiplicative error terms can be incorporated in RUM models. Fosgerau and Bierlaire (2009) find that in most cases a multiplicative specification outperforms an additive specification and that the improvement can sometimes even be larger than that gained from allowing for unobserved heterogeneity.

Hess et al. (2008) have provided some evidence on asymmetrical responses to increases and decreases in the levels of the attributes describing alternatives in DCE. Their results seem to support the argument that utility functions depend on changes in the values of attributes rather than the actual values themselves. Allowing for asymmetrical preferences among respondents means that a larger share of the behavioural pattern can be explained in a deterministic manner rather than using MXL model specifications.

6.2. Experimental design theory

Interest in the experimental design of DCE has increased in environmental economics, although the main research has been undertaken in the field of marketing and transportation. There is a growing attention on experimental design affecting an important factor in discrete choice analysis as the model specification used to analyse the data obtained.[12] The empirical applications in the field of environmental economics have mainly relied on the use of orthogonal designs (Louviere et al., 2000). However, the use of orthogonal designs in DCE has been recently challenged (Huber and Zwerina, 1996; Kessels et al., 2006). A recent trend in the literature has started to move away from orthogonal designs towards designs which relate to the econometric models used in fitting DCE data.

Huber and Zwerina (1996) were first to relate the statistical properties of the experimental design to the estimated econometric model, showing that targeting the minimum asymptotic standard errors of the parameter estimates rather than orthogonality in the design resulted in more efficient designs that improved the reliability of the parameter estimates. As argued by Train (2003), non-linear models, such as DCE, should rather be concerned about the differences in the utility functions. Efficient or optimal designs attempt to link the experimental design generation process with the smaller asymptotic standard errors of the parameter estimates based on the idea that the concern in DCE it is not the correlation structure between the attributes but the correlations of the differences in the attributes. Orthogonality, within the context of DCE, should relate towards the correlation between the differences in the utility functions of the alternatives within the data rather than the actual values observed for each attribute (Bliemer and Rose, 2006).

[12] Louviere (2006), for example, argues that researchers should provide the level of efficiency of their designs and make it available for peer-reviewing.

The central argument against the use of orthogonal designs in the context of DCE is that the statistical properties of orthogonal designs do not hold. Orthogonality in linear models not only avoids problems of multicollinearity in the estimated model, but more importantly, it minimises the standard errors of the parameter estimates. This is true because orthogonality minimises the elements of the model's variance-covariance matrix. However, in non-linear models such as the models used in DCE, orthogonality has lower relation with the expected AVC matrix given that, on the one hand, the log-likelihood function and second derivatives of discrete choice models are dependent on the choice probabilities obtained for the choice data and, on the other hand, the only thing that matter in DCE is the differences in the utility functions between the chosen and non-chosen alternatives (Bliemer et al., 2009).

Given that the parameters estimation is based on data sets obtained in DCE and not from the design itself, orthogonality will likely to be lost in the estimation process. Some reasons for losing orthogonality are missing responses, blocks not equally represented in the data set, inclusion of sociodemographic variables in the estimated models, or attribute levels not being equidistant in spacing. So, in practice, an orthogonal design is likely to end up in non-orthogonal data (Bliemer and Rose, 2006).

The recent development of optimal or efficient designs has served to question the need for orthogonality. These designs focus on generating parameter estimates with the lowest standard errors by determining the AVC matrix of the parameters.[13] In this case, prior information on the parameter estimates (based for example on the results of similar studies or on a pilot survey) is used for deriving the AVC matrix. The efficient design will attempt to combine the choice situations so that the standard errors of the parameter estimates are minimised. In other words, the objective would be to minimise the efficiency error, as it has been shown in Section 3.2.

[13] The square roots of the diagonal elements of the AVC matrix are the asymptotic standard errors.

Advanced designs are special efficient designs that either relax some assumptions or impose further constraints for practical reasons (Bliemer and Rose, 2006): constrained designs are those in which some combinations of the attribute levels are imposed so that infeasible combinations are avoided; pivot designs are those in which the assumption that all respondents face the same choice situations are relaxed in order to create more realistic choice situations for a particular respondent; and designs with covariates are those in which individuals' characteristics are included in the model in order to optimise the design for each group of respondents.

From the above discussion, it can be concluded that efficient designs will outperform orthogonal designs if any information about the parameters is available. However, efficient designs create their own challenges: firstly, the determination of the priors used for generating the design, and secondly, that the final model form should be known in advance. Information on priors may be obtained from similar studies in the literature or by conducting a pilot study. Furthermore, one could argue that at least we could know that some parameters should have a negative sign (for example the cost attribute) or that desirable attributes such as a protected landscape should have a positive sign. The advantage of efficient designs is that this prior information can help developing experimental designs where either parameter estimates have lower standard errors or the sample size required is smaller. Some authors have been concerned about the impact of prior parameter estimates on the final model results. However, Blimer et al. (2009) argue that "misspecification of priors may decrease the efficiency of the design but the efficiency will in general still be better than assuming zero priors."

The second challenge in generating efficient designs is that efficiency is related to the econometric model that will most likely be used for estimating the parameters after collecting the data. This is due to the fact that the efficiency of a DCE is related to the AVC matrix of the model to be estimated and that different econometric models will have a different AVC matrix. So, depending on the estimated model, the same experimental design will end up in different levels of efficiency. Bliemer et al. (2009) find that the efficiency of the design is maximized only when the model assumed in generating the design is the model that is fitted during estimation. This issue may raise concerns about how to generate efficient designs under uncertainty over the final model that will be estimated. Rose et al. (2009) propose a design generation based on a model averaging approach over different econometric models. Future research should concentrate on linking the statistical properties of DCE designs and respondent's behaviour with, for example, information processing strategies (Bliemer and Rose, 2006).[14]

6.3. Model estimation

As it has been shown is Section 4.1, three inter-related issues are involved in the specification of heterogeneity in MXL models: the selection of the parameters that should be random, the choice of mixing distribution for these parameters and their economic interpretation. The selection of random parameters was already discussed in Section 4.1. This Section will concentrate on the problems that may arise when choosing mixing distributions for random parameters and its consequences in terms of allowing positive and negative coefficient values. This is an important yet unsolved issue because an inappropriate choice of mixing distribution is a potential source of model misspecification. It is also important to bear in mind that model fit may not always be an appropriate indicator of model performance.

Estimating WTP measures using MXL models may be complicated because of the difficulty of maintaining theoretical consistency and actual behaviour of decision makers, constrained by the data collection and model specification. It is not infrequent to find applications of DCE using MXL models that have fixed the price coefficient, especially when the researcher aims to estimate the distribution of individuals' WTP for environmental attributes (see for example

[14] The Journal of Choice Modelling is currently (summer 2009) preparing a special issue dedicated to experimental design.

Layton and Brown (2000) or Revelt and Train (1998)). The main problem of fixing the distribution of the price coefficient when in fact scale varies over observations is that variation in scale would be erroneously translated into variation in WTP (Train and Weeks, 2005).

Another common strategy is to find a 'reasonable' distribution of the price coefficient acknowledging, on the one hand, that it should be chosen according to the data generating process and, on the other, that it should be consistent with theories of rational economic behaviour. In other words, obtaining a positive price coefficient (i.e. a positive marginal utility of income) seems inconsistent with the rational economic behaviour underlying the theory of random utility maximisation.[15] It is also important to bear in mind that if the environmental attribute and price coefficients are random, the WTP will follow an unknown distribution that will need to be simulated. As a consequence, the researcher may be in many cases interested in obtaining so-called reasonable WTP distributions (i.e. preventing WTP for changing sign or being unreasonably large).

The Normal distribution has been commonly used in MXL models. Given that this distribution is symmetric and unbounded, the researcher assumes that positive and negative values for this coefficient can be found in the population. However, if the true distribution yielded strictly negative values with a mean close to zero and a long tail into the negative side, one cannot decide whether a non-zero probability of a positive coefficient is revealed by the data or if it is an artifact of the symmetrical nature of the Normal distribution. So, the Normal distribution may be used for parameters attributes where the researcher has no a priori assumption about

[15] If all correlated factors, such as individual's prestige effects, were properly accounted for, the marginal utility of income should be negative and it would be inappropriate to use unbounded distributions of the price coefficient. However, the use of distributions with flexible bounds may allow for these correlated factors not accounted for to reveal themselves in particular empirical applications (Hess et al., 2005).

its sign but it could lead to false conclusions indicating a probability of sign change in the random coefficient that does not exist in the population.

The most common distribution for coefficients with an explicit sign assumption is the Lognormal distribution. In this case, the distribution is bounded at zero so that there is no sign change. However, this distribution has a long tail on the unbounded side so that extremely large values may be found and it can overestimate standard deviations. Other distributions with a fixed bound include the Gamma, Rayleigh or Exponential distributions. As opposite to the previous case, the problem with distributions with a fixed bound at zero is that they will not allow for a sign change due to values revealed by the data, so the fit can be poorer.

In sum, while using an unbounded distribution the analyst may be forced to explain a significant probability of a sign change by alleging that some agents have 'irrational' preferences, by arbitrarily constraining the model to only produce the sign coefficient requirement the analyst may ignore the impact of data or model imperfections. As a consequence, Hess et al. (2005) argue that the best option would be to use bounded distributions where the bounds are directly estimated from the data. This is the case of the Triangular distribution that not only avoids the long tails of the Normal distribution or the strict bounds of the Lognormal distribution but it can be asymmetrically defined. Another flexible distribution is the Johnson S_b distribution. Despite the need to estimate four parameters, this distribution has been proved to provide good results by Train and Sonnier (2005) and Hoyos et al. (2009). Other possible approaches are the use of empirical distributions reflecting the actual distribution found in the sample (Hess et al., 2005) or to use censored distributions (Train and Sonnier, 2005).

The above discussion highlights the importance of assuming mixing distributions that are theoretically consistent but that allow, at the same time, for model imperfections. By doing this, the analyst may distinguish whether her findings are inconsistencies with economic theory or econometric artifacts due to the complexity of specifying taste heterogeneity in discrete choice models (Hess et al., 2005).

Another related way of dealing with the problem of estimating WTP measures is to look at the consequences of placing the distributional assumptions in the preference space or in the willingness-to-pay space (Train and Weeks, 2005). In the first case, a distribution of coefficients in the utility function is specified and the distribution of the WTP is later derived. In the second case, a distribution of WTP is directly specified and the distribution of coefficients is later derived. Train and Weeks (2005) find that models in preference space fit better the data at the cost of less reasonable distributions of WTP compared to those models that work directly in the WTP space, where the number of individuals having untenably large WTP was significantly lower.

There has been a recent interest in modelling the choices of individuals under the so-called *individual-level* discrete choice models. By differentiating the distribution of tastes at the individual level, policy-makers may be provided with valuable information. Furthermore, by estimating models for single individuals empirical distributions of preferences can be directly estimated so that random parameters of finite mixture models would not be necessary (Louviere, 2006). However, the problem arises as how to obtain sufficient choice observations to estimate models for individuals. Two approaches have emerged: top-down and bottom-up approach. Top-down approaches estimate individual-level parameters indirectly, so that the aggregate preference distribution estimated in MXL models is combined with information of individuals' choices to calculate conditional estimates of respondents' preferences. Revelt and Train's (2001) "maximum likelihood with conditioning of individual tastes" approach derives the distribution of tastes of an individual conditional on the observed choices of that particular individual and the estimates of the distribution of tastes in the population. A hierarchical Bayes (HB) estimation approach has been shown to provide similar results (Huber and Train, 2001). The second approach proposed by Louviere et al. (2009) combines optimal designs with repeated best-worst choice questions in order to directly estimate individual parameters. In theory, if the mixing distributions are correctly specified and the number of choices per person sufficiently large, both approaches should converge. However, if the assumptions about the sample distribution of preferences are incorrect, inference based on the top-down approach will be biased and incorrect (Louviere et al., 2009).

Another recent focus of concern among academics refers to the problem of endogeneity in DCE. An explanatory variable is said to be endogenous when it is correlated (i.e. not independent) of the unobserved factors. Endogeneity in DCE may appear, for example, because of the influence of unobserved attributes or because the choices of decision-makers may be inter-related (Train, 2003). Several methods have been proposed to estimate DCE in the presence of endogenous explanatory variables, such as the Berry-Levinsohn-Pakes approach (Berry et al., 1995), the control function approach (Petrin and Train, 2009) or the full maximum likelihood approach (Park and Gupta, 2009). Louviere et al. (2005) describe recent progress in dealing with endogeneity in DCE.

Finally, the increasing difficulty in maximising more complex log-likelihood functions has motivated the development of new algorithms, such as the Expectation-Maximisation (EM) algorithm. The procedure to estimate models using the EM algorithm is explained in Train (2003).

6.4. Hypothesis testing

Recent developments in respect to hypothesis testing within the DCE framework have attempted to fill some gaps. In respect to nonlinear specifications (piecewise linear specifications, power series expansions and Box-Cox transformations), they can be easily tested using LR tests. However, other potentially problematic issues such as model misspecification or the appropriateness of the distributional assumptions in RPL models are increasingly being analysed.

Model misspecification generally invalids statistical inference, although it is rarely tested in DC models. For this purpose, Fosgerau (2008) has recently proposed the use of a nonparametric test of functional form, the Zheng test, to discrete choice models. The appropriateness of distributional assumptions of the random parameters included in RPL models is also rarely tested. For this purpose, Fosgerau and Bierlaire (2007) have proposed a test based on seminonparametric (SNP) techniques.

7. Conclusions

Environmental valuation with DCE is playing an increasingly significant role in environmental decision-making. By turning the focus onto how preferences for nonmarket goods are organised, DCE aim at identifying the utility that individuals derive from the attributes conforming an environmental good or service under valuation. Despite the complexity of the task and the potential loss of incentive compatibility, the move from the CVM to the DCE has brought many potential advantages in terms of its multi-attribute approach and its ability to estimate marginal value of changes in the levels of those attributes. Another advantage encountered using DCE is that the incidence of ethical protesting seems to be lower (Hanley et al., 2001). DCE can also provide the opportunity to elicit a deeper understanding of the trade-offs between different attributes (Adamowicz et al., 1998; Jin et al., 2006b). Further, when valuing multi-attribute programmes, DCE can be significantly cheaper to implement because it requires only one single questionnaire. Hence, given its inherent flexibility DCE may be, in many circumstances, a more useful tool compared to the CVM. In the words of Hanley et al. (2001), “DCE seem to be ideally suited to inform the choice and design of multidimensional policies”.

From the early work mainly focused on recreation demand, the use of DCE has rapidly increased in the fields of environmental and ecological economics as it can be shown by the vast amount of applications presented in specialised congresses. However, given its infancy state, more research is still needed in order to determine which experiments and response procedures are more likely to produce more reliable estimates. As many authors have highlighted, it is clear that in order to be well established DCE need more testing of its properties, as it has been the case with CVM (Bateman et al., 2008).

Environmental valuation using SP methods is complex because it requires knowledge of many different disciplines such as economic theory, experimental and survey design theory, data collection and econometric analysis. The research in environmental and ecological economics is further complemented by the developments in psychology, marketing, transportation, decision theory or statistics. In fact, interdisciplinary collaboration may be seen as a major contributor to this field. As Louviere (2006) argues: "virtually all major scientific breakthroughs result from cross-disciplinary innovations".

The research on environmental valuation using DCE has been intense in the last fifteen years but it can be still gathered as an infant research field. Many challenges need to be overcome. One of the main issues surrounding DCE is that of choice task complexity and cognitive effort. This may be especially true when respondents are asked to trade between complex and unfamiliar goods and services such as those generally involved in environmental valuation. Undertaking a DCE requires firstly a sound construction of the contextual frame in which choice occasions occur. Experimental economics can significantly contribute to the SP literature at this stage by better understanding how individuals confront institutions, incentives and information delivered in surveys and so improve the survey design (Shogren, 2006). Experimental design has been also shown to have a major role in environmental valuation. Existing research has shown that more attention should be put to the choice set construction since changes in the design may systematically affect the parameter estimates and the variances of the error terms (Adamowicz and Deshazo, 2006). Louviere (2006) warns about the impossibility of learning about individuals' decision rules unless full factorial designs are used. At the estimation stage, differences in the preferences of individuals (i.e. preference heterogeneity) should be adequately identified in order to obtain unbiased estimates of demand. Endogeneity also represents a major future challenge of environmental valuation. Finally, the researcher needs to embrace model uncertainty by choosing the 'true' model. In the absence of a single dominant selection criterion, some authors have suggested to obtain more preference information by pooling data and use model averaging approaches in order to obtain more efficient estimates (Layton and Lee, 2006). Rather than generating in more complex statistically fitting choice models, some authors have argued for developing a better behavioural theory guiding both model specification and empirical research (Louviere, 2006).

Acknowledgments

The author acknowledges the financial support from the Department of Environment of the Basque Government through IHOBE, S.A. and from the Department of Education of the Basque Government through grant IT-334-07 (UPV/EHU Econometrics Research Group).

References

Adamowicz W., P. Boxall, M. Williams, and J. Louviere, 1998. Stated preference approaches for measuring passive use values: choice experiment and contingent valuation. American Journal of Agricultural Economics, 80:64-75.

Adamowicz W. and J. R. Deshazo, 2006. Frontiers in stated preference methods: an introduction. Environmental and Resource Economics, 34:1-6.

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.

Adamowicz W., J. Swait, P. Boxall, J. Louviere, and M. Williams, 1997. Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. Journal of Environmental Economics and Management, 32:65-84.

Alpizar F., F. Carlsson, and P. Martinsson, 2001. Using choice experiments for non-market valuation. Economic Issues, 8:83-110.

Araña J. E. and C. J. Leon, 2009. Understanding the use of non-compensatory decision rules in discrete choice experiments: The role of emotions. Ecological Economics, 68:2316-2326.

Bateman I. J., R. T. Carson, B. H. Day, D. Dupont, J. Louviere, S. Morimoto, R. Scarpa, and P. Wang, 2008. Choice set awareness and ordering effects in choice experiments. Paper presented to the 2008 EAERE Conference in Gothenburg, Sweden.

Berry S., J. Levinsohn, and A. Pakes, 1995. Automobile prices in market equilibrium. Econometrica, 63:841-890.

Bienabe E. and R. R. Hearne, 2006. Public preferences for biodiversity conservation and scenic beauty within a framework of environmental services payments. Forest Policy and Economics, 9:335-348.

Birol E., K. Karousakis, and P. Koundouri, 2006. Using a choice experiment to account for preference heterogeneity in wetland attributes: The case of Cheimaditida wetland in Greece. Ecological Economics, 60:145-156.

Bliemer M. C. J.and J. M. Rose, 2006. Designing stated choice experiments: state-of-the-art. International Conference on Travel Behaviour Research. Kyoto, 16-20 August, 2006.

Bliemer M. C. J., J. M. Rose, and D. A. Hensher, 2009. Efficient stated choice experiments for estimating nested logit models. Transportation Research Part B: Methodological, 43:19-35.

Brey R., P. Riera, and J. Mogas, 2007. Estimation of forest values using choice modeling: an application to Spanish forests. Ecological Economics, 64:305-312.

Brownstone D., 2001. Discrete choice modeling for transportation. In: DA Henser (Editor), Travel behaviour research: the leading edge. Pergamon, Amsterdam, pp. 97-124.

Campbell D., W. G. Hutchinson, and R. Scarpa, 2008. Incorporating discontinous preferences into the analysis of discrete choice experiments. Environmental and Resource Economics, 41:401-417.

Caparros A., J. L. Oviedo, and P. Campos, 2008. Would you choose your preferred option? Comparing choice and recoded ranking experiments. American Journal of Agricultural Economics, 90:843-855.

Carlsson F. and P. Martinsson, 2001a. Do hypothetical and actual marginal willingness to pay differ in choice experiments?: application to the valuation of the environment. Journal of Environmental Economics and Management, 41:179-192.

Carlsson F. and P. Martinsson, 2001b. Do hypothetical and actual marginal willingness to pay differ in choice experiments?: application to the valuation of the environment. Journal of Environmental Economics and Management, 41:179-192.

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, 5:351-368.

Cox, D., 1961. Test of separate families of hypotheses. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1. Berkeley: University of California Press.

Cox, D., 1962. Further results on tests of separate families of hypotheses. Journal of the Royal Statistical Society, Series B, 24, 406-424.

Davidson, R. and J. MacKinnon, 1981. Several tests for model specification in the presence of alternative hypotheses. Econometrica, 49, 781-793.

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, 44:123-143.

Fosgerau M., 2008. Specification testing of discrete choice models: a note on the use of a nonparametric test. Journal of Choice Modelling, 1:26-39.

Fosgerau M. and M. Bierlaire, 2009. Discrete choice models with multiplicative error terms. Transportation Research Part B: Methodological, 43:494-505.

Fosgerau M. and M. Bierlaire, 2007. A practical test for the choice of mixing distribution in discrete choice models. Transportation Research Part B: Methodological, 41:784-794.

Foster V. and S. Mourato, 2003. Elicitation format and sensitivity to scope. Environmental and Resource Economics, 24:141-160.

Goldberg I. and J. Roosen, 2007. Scope insensitivity in health reduction studies: a comparison of choice experiments and the contingent valuation method for valuing safer food. Journal of Risk and Uncertainty, 34:123-144.

Greene, W.H., 2003. Econometric Analysis (5th edition). Pearson Education, New York.

Hanemann W. M., 1984. Discrete/continuous models of consumer demand. Econometrica, 52:541-561.

Hanemann W. M.and B. J. Kanninen, 1999. The statistical analysis of discrete-response CV data. In: IJ Bateman and KG Willis (Editors), Valuing environmental preferences. Oxford University Press, Oxford.

Hanley N., S. Mourato, and R. E. Wright, 2001. Choice modelling approaches: a superior alternative for environmental valuation? Journal of Economic Surveys, 15:435-462.

Hanley N., W. Adamowicz, and R. E. Wright, 2005. Price vector effects in choice experiments: an empirical test. Resource and Energy Economics, 27:227-234.

Hausman J. and D. McFadden, 1984. Specification tests for the multinomial logit model. Econometrica, 52:1219-1240.

Henser D. A., J. M. Rose, and W. H. Greene, 2005. Applied choice analysis. A primer, Cambridge University Press, New York.

Hensher D. A., 2007. Attribute processing in choice experiments and implications on willingness to pay. In: BJ Kanninen (Editor), Valuing environmental amenities using stated choice studies. A common sense approach to theory and practice. Springer, The Netherlands, pp. 135-158.

Hensher D. A., 2008. Joint estimation of process and outcome in choice experiments and implications for willingness to pay. Journal of Transport Economics and Policy, 42:297-322.

Hensher D. A., J. M. Rose, and W. H. Greene, 2005. Applied Choice Analysis. A Primer, Cambridge University Press, New York.

Hess S., M. Bierlaire, and J. W. Polak, 2005. Estimation of value of travel-time savings using mixed logit models. Transportation Research Part A: Policy and Practice, 39:221-236.

Hess S., J. M. Rose, and D. A. Hensher, 2008. Asymmetric preference formation in willingness to pay estimates in discrete choice models. Transportation Research Part E: Logistics and Transportation Review, 44:847-863.

Horne P., P. C. Boxall, and W. L. Adamowicz, 2005. Multiple-use management of forest recreation sites: a spatially explicit choice experiment. Forest Ecology and Management, 207:189-199.

Hoyos D., P. Mariel, and J. Fernandez-Macho, 2009. The influence of cultural identity on the WTP to protect natural resources: Some empirical evidence. Ecological Economics, 68:2372-2381.

Huber J. and K. Train, 2001. On the similarity of classical and Bayesian estimates of individual mean partworths. Marketing Letters, 12:259-269.

Huber J. and K. Zwerina, 1996. The importance of utility balance in efficient choice designs. Journal of Marketing Research, 33:307-317.

Jin J., Z. Wang, and S. Ran, 2006b. Comparison of contingent valuation and choice experiment in solid waste management programs in Macao. Ecological Economics, 57:430-441.

Jin J., Z. Wang, and S. Ran, 2006a. Comparison of contingent valuation and choice experiment in solid waste management programs in Macao. Ecological Economics, 57:430-441.

Johnson F. R. and K. E. Mathews, 2001. Sources and effects of utility-theoretic inconsistency in Stated Preference surveys. American Journal of Agricultural Economics, 83:1328-1333.

Kahneman D.and S. Frederick, 2002. Representativeness revisited: attribute substitution in intuitive judgment. In: T Gilovich, D Griffin, and D Kahneman (Editors), Heuristics and biases: the psychology of intuitive judgment. Cambridge University Press, New York.

Kessels R., P. Goos, and M. Vandebroek, 2006. A comparison of criteria to design efficient choice experiments. Journal of Marketing Research, 43:409-419.

Krinsky I. and A. Robb, 1986. On approximating the statistical properties of elasticities. The Review of Economics and Statistics, 68:715-719.

Ladenburg J. and S. r. B. y. Olsen, 2008. Gender-specific starting point bias in choice experiments: Evidence from an empirical study. Journal of Environmental Economics and Management, 56:275-285.

Lancaster K., 1966. A new approach to consumer theory. Journal of Political Economy, 74:132-157.

Lancsar E. and J. Louviere, 2006. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? Health Economics, 15:797-811.

Lancsar E. and J. Louviere, 2008. Conducting discrete choice experiments to inform healthcare decision making. A user's guide. Pharmaeconomics, 26:661-677.

Lanz, B., A. Provins et al., 2009. Investigating willingness to pay - willingness to accept asymmetry in choice experiments. Conference Proceedings.

Layton D. F. and G. Brown, 2000. Heterogeneous Preferences Regarding Global Climate Change. The Review of Economics and Statistics, 82:616-624.

Layton D. F. and S. T. Lee, 2006. Embracing model uncertainty: strategies for response pooling and model averaging. Environmental and Resource Economics, 34:51-85.

Louviere J., 2006. What you don't know might hurt you: some unresolved issues in the design and analysis of discrete choice experiments. Environmental and Resource Economics, 34:173-188.

Louviere J. and D. A. Henser, 1982. On the design and analysis of simulated choice or allocation experiments in travel choice modelling. Transportation Research Record, 890:11-17.

Louviere J., D. A. Hensher, and J. Swait, 2000. Stated choice methods: analysis and applications, Cambridge University Press, Cambridge.

Louviere J., D. J. Street, L. Burgess, N. Wasi, T. Islam, and A. Marley, 2009. Modeling the choices of individual decision-makers by combining efficient choice experiment design with extra preference information. Journal of Choice Modelling, 1:128-163.

Louviere J., K. Train, M. Ben, C. Bhat, D. Brownstone, T. A. Cameron, R. T. Carson, J. R. Deshazo, D. Fiebig, W. Greene, D. Hensher, and D. Waldman, 2005. Recent progress on endogeneity in choice modeling. Marketing Letters, 16:255.

Louviere J. and G. Woodworth, 1983. Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. Journal of Marketing Research, 20:350-367.

Lusk J. L. and T. C. Schroeder, 2004. Are choice experiments incentive compatible? A test with quality differentiated beef steaks. American Journal of Agricultural Economics, 86:467-482.

McFadden D., 1974. Conditional logit analysis of qualitative choice behaviour. In: P Zarembka (Editor), Frontiers in Econometrics. Academic Press, New York.

McFadden D., 1995. Computing willingness-to-pay in random utility models.

McFadden D. and K. Train, 2000. Mixed MNL models for discrete response. Journal of Applied Econometrics, 15:447-470.

McIntosh E. and M. Ryan, 2002. Using discrete choice experiments to derive welfare estimates for the provision of elective surgery: Implications of discontinuous preferences. Journal of Economic Psychology, 23:367-382.

Meyerhoff J.and U. Liebe, 2009. Discontinuos preferences in choice experiments: evidence at the choice task level. XVII EAERE Conference. Amsterdam 2009.
Ref Type: Conference Proceeding

Mogas J., P. Riera, and J. Bennett, 2006. A comparison of contingent valuation and choice modelling with second-order interactions. Journal of Forest Economics, 12:5-30.

Morey E., R. Rowe, and M. Watson, 1993. A repeated nested-logit model of atlantic salmon fishing. American Journal of Agricultural Economics, 75:578-592.

Morrison M. D., J. W. Bennett, R. K. Blamey, and J. Louviere, 2002. Choice modelling and tests of benefit transfer. American Journal of Agricultural Economics, 84:161-170.

Nielsen A. B., S. B. Olsen, and T. Lundhede, 2007. An economic valuation of the recreational benefits associated with nature-based forest management practices. Landscape and Urban Planning, 80:63-71.

Park S. and S. Gupta, 2009. A simulated maximum likelihood estimator for the random coefficient logit model using aggregate data. Journal of Marketing Research, forthcomming.

Payne J. W., J. R. Bettman, and E. J. Johnson, 1993. The adaptive decision maker, Cambridge University Press, Cambridge.

Petrin A. and K. Train, 2009. A control function apprach to endogeneity in consumer choice models. Journal of Marketing Research, forthcomming.

Puckett S. M. and D. A. Hensher, 2009. Revealing the extent of process heterogeneity in choice analysis: An empirical assessment. Transportation Research Part A: Policy and Practice, 43:117-126.

Revelt D. and K. Train, 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. The Review of Economics and Statistics, 80:647-657.

Revelt D.and K. Train, 2001. Customer-specific taste parameters and mixed logit: households' choice of electricity supplier. University of Berkeley.

Rolfe J., J. Bennett, and J. Louviere, 2000. Choice modelling and its potential aplication to tropical rainforest preservation. Ecological Economics, 35:289-302.

Rose J. M., R. Scarpa, and M. C. J. Bliemer, 2009. Incorporating model uncertainty into the generation of efficient stated choice experiments: a model averaging approach. International Choice Modelling Conference. Harrogate, 30 march - 1 april, 2009.
Ref Type: Conference Proceeding

Samuelson W. and R. Zeckhauser, 1988. Status quo bias and decision making. Journal of Risk and Uncertainty, 1:7-59.

Scarpa R., D. Campbell, and W. G. Hutchinson, 2007. Benefit estimates for landscape improvements: sequential bayesian design and respondents' rationality in a choice experiment. Land Economics, 83:617-634.

Schkade D. A. and J. W. Payne, 1994. How people respond to contingent valuation questions: a verbal protocol analysis of willingness to pay for an environmental regulation. Journal of Environmental Economics and Management, 26:88-109.

Scott A., 2002. Identifying and analysing dominant preferences in discrete choice experiments: an application in health care. Journal of Economic Psychology, 23:383-398.

Shogren J. F., 2006. Valuation in the lab. Environmental and Resource Economics, 34:163-172.

Street D. J., L. Burgess, and J. J. Louviere, 2005. Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments. International Journal of Research in Marketing, 22:459-470.

Train K., 2003. Discrete Choice Methods with Simulation, Cambridge University Press, New York.

Train K.and G. Sonnier, 2005. Mixed logit with bounded distributions of correlated partworths. In: R Scarpa and A Alberini (Editors), Applications of simulation methods in Environmental and Resource Economics. Springer, Dordrecht, The Netherlands, pp. 1-16.

Train K.and M. Weeks, 2005. Discrete choice models in preference space and willingness-to-pay space. In: R Scarpa and A Alberini (Editors), Applications of simulation methods in Environmental and Resource Economics. Springer, Dordrecht, The Netherlands, pp. 1-16.

Tversky A., 1972. Elimination by aspects: a theory of choice. Psychological Review, 79**:**281-299.

Xu W., B. R. Lippke, and J. Perez-Garcia, 2003. Valuing biodiversity, aesthetics and job losses associated with ecosystem management using stated preferences. Forest Science, 49**:**247-257.