

APPLICATION OF CHOICE EXPERIMENT: THEORETICAL ASPECT

RESEARCH ARTICLES

Sri Lanka Journal of
Economic Research
Volume 3 (2)
December 2015: 31-54
Sri Lanka Forum of
University Economists

SLJER

Wasantha Athukorala
Muditha Karunarathna



Abstract

Among the environmental valuation methods, the Choice Experiment (CE) method is considered to be the most appropriate method for valuing benefits of attributes related to a particular environmental commodity. This is because of the CE method allows not only for estimation of the value of the environmental good as a whole, but also for the implicit values of its attributes. Under this method a sample of people is asked to choose their most preferred alternatives from a sequence of grouped options that relate to different management strategies. Each option is described in terms of its outcomes and a personal monetary cost to be borne personally by the respondent. In analysing the choices made by respondents, it is possible to infer the trade-off that people are willing to make between money and greater benefits of different attributes. This paper aims at explaining the basic steps of undertaking a choice experiment study which is increasingly becoming popular technique in both the developed as well as in developing countries. Researchers who are interested in applying CE method for their research can use this as a basic guidance for their work.

Keywords: *Environmental valuation, Choice Experiment (CE) method, Management strategies, Developing countries.*

Wasantha Athukorala

Senior Lecturer, Department of Economics and Statistics, University of Peradeniya, Sri Lanka. Email: wathukorala@yahoo.com

Muditha Karunarathna

Senior Lecturer, Department of Economics and Statistics, University of Peradeniya, Sri Lanka.

INTRODUCTION

The valuation of nonmarket goods is one of the principle issues addressed by environmental economics research (Champ et al., 2004). When competitive markets exist, market prices are the appropriate measure of social well-beings. However, in practice, all markets do not function exactly in the manner assumed by economic theory. In such cases market prices are not the best available approximate measures of social values of goods and services (Freeman, 2003). There are two main valuation techniques which are widely used in environmental economics. The first method is the revealed preference techniques where people's preferences for particular commodity are revealed through their actions in related markets. The second method is the stated preference techniques that require people to state the strength of their preferences and hence reveal the values they enjoy through structured questionnaires (Bishop and Romano, 1998). This method does not involve any reliance on market data.

For market based valuation techniques, the benefit generated by the environmental commodities must be bought and sold in markets. Hence, the techniques are mostly suitable for application where direct use benefits are involved. As both consumer and producer receive the benefits, consumer surplus and producer surplus can be used to measure the total benefits received from use value of the commodity. Therefore, it is clear that if there are sufficient observations of trade, it is possible to use standard economic techniques to estimate values for both buyers and sellers (Freeman, 2003). For example, if a species is under threat of extinction, the cost of a captive breeding program may be used to estimate the benefit being provided by its continued survival. Another approach involves the estimation of how much it would cost to replace the lost of a forest area with a substitute. This replacement cost technique is widely used in various analyses because of its reliability as well as the simplicity of capturing the relevant cost.

Limitations of the market based or revealed preference techniques, led to the development of stated preference techniques (Champ et al., 2004). In this type of technique, a sample of people are asked about their preferences for a sensitive asset under a hypothetical set of circumstances. A number of different methods have been developed to inquire about peoples' preferences. The first stated preference technique

to be developed was the contingent valuation method (CVM)¹. Originally, this method required that a sample of people be asked the amount they would be willing to pay to secure an improvement in a particular aspect of the environmental commodity. More recently, this technique has been refined to accommodate a dichotomous choice version that involves people being asked if they would or would not support a proposal to improve the existing system given some personal monetary cost (Karunaratna, 2012). This is the basic idea of choice experiment method. In a CE, individuals are presented with a choice set or series of choice sets that are framed with various attributes and attribute levels and are asked to choose one bundle at a varied set of price and attribute levels. Consumers' willingness to accept (WTA) compensation payment for each attribute is then computed from estimates of econometric models.

Although CE method is the most reliable approach to estimate the non-market benefits, application in developing countries is limited. Given its complexity as well as the requirement of the theoretical and empirical knowledge, most people are reluctant to apply this methodology in their studies. This paper attempts to explain the basic steps of the CE procedure which will enhance theoretical as well as empirical knowledge in this area. Basic methodology has a theoretical grounding in Lancaster's attribute theory of consumer choice (Lancaster, 1966) and an econometric basis in models of random utility (Luce, 1959; McFadden, 1974). Therefore, RUM is explained in the next section.

RANDOM UTILITY MODELS (RUM)

The CE methods rely on the random utility model framework to provide a utility theoretical interpretation of the discrete responses observed from the respondents. Garber-Yonts (2001) provided the basic steps of RUM and a derivation of WTP compensation that is explained below. Given a set of alternatives A_n , presented to an individual n , the probability that any one alternative i is chosen is given by:

$$P(i / A_n) = \Pr(U_{in} \geq U_{jn}, V_j \in A_n) \quad (1)$$

¹ The idea of CVM was first suggested by Ciriacy-Wantrup (1947), and the first study ever done was in 1961 by Davis (1963).

Where, U_{in} is the utility that individual n achieves by choosing alternative i . According to the random utility theory, the utility which is not directly observable can be partitioned into a deterministic component and a random component (Ben-Akiva and Lerman 1985; Garber-Yonts, 2001). The accompanying assumption is that the individual knows their utility function with certainty, however with other measurement errors, utility can be stochastic:

$$U_{in} = V_{in} + \varepsilon_{in} \quad (2)$$

Where, V_{in} is the mean and the random disturbance of the stochastic random utility function. The specification of V_{in} includes a vector of attribute of alternative i , X_{in} , which includes a price or bid variable, and a vector of characteristics of the respondent, H_n , including income (Garber-Yonts, 2001). Thus model can be written as Equation 3:

$$U_{in} = \beta' f(X_{in}, H_n) + \varepsilon_{in} \quad (3)$$

Where, the deterministic component is here specified as linear in parameters, though the function $f(.)$ can be nonlinear. However, when choosing the functional form, there is a trade-off between the benefits of assuming a less restrictive formulation and the complications that arise from doing so. This is especially relevant for the way income enters the utility function (Garber-Yonts, 2001). A simpler functional form (e.g. linear in income) makes estimation of the parameters and calculation of welfare effects easier, but the estimates are based on restrictive assumptions (Ben-Akiva and Lerman, 1985). Most often researchers have been inclined to use a simpler linear in the parameters utility function. Another important thing is that the error term enters the utility function as an additive term. This assumption, although restrictive, greatly simplifies the computation of the results and the estimation of welfare measures. With the indirect utility specified as above, the individual seeks to maximize utility such that:

$$P_n(i/A_n) = P(\beta' f(X_{in}, H_n) + \varepsilon_{in} > \beta' f(X_{jn}, H_n) + \varepsilon_{jn})$$

$$P_n(i/A_n) = P(\beta' f(X_{in}, H_n) - \beta' f(X_{jn}, H_n) \geq (\varepsilon_{jn} + \varepsilon_{in})); i, j \in A_n, i \neq j \quad (4)$$

It becomes clear that unless H_n enters the function $f(\cdot)$ nonadditively, it appears identically on both sides of the inequality and cancels out of the function. Thus, H_n must enter nonadditively if the effects of respondent characteristics on choice are to be measured (Garber-Yonts, 2001). If ε_{in} and ε_{jn} are assumed to be extreme value independently and identically distributed (IID) with scale parameter μ , then $\varepsilon^* = \varepsilon_{jn} - \varepsilon_{in}$ is logistically distributed (Ben-Akiva and Lerman, 1985). This distributional assumption approximates the normal distribution which leads to the multinomial logit (MNL) model for the choice probabilities (McFadden, 1974; Ben-Akiva and Lerman, 1985). This is the simplest version of the analysis of multinomial outcomes when comparing with conditional logit (CL) model and RPL model. MNL model can be given as Equation 5:

$$P_n(i/A_n) = e^{\mu^{V_{jn}}} / \sum_{j \in A_n} e^{\mu^{V_{jn}}} = e^{\mu \beta' f(X_{in}, H_n)} / \sum_{j \in A_n} e^{\mu \beta' f(X_{jn}, H_n)} = e^{\mu \beta' f(X_{in}, H_n)} \quad (5)$$

Since μ appears as a multiplicative constant on every parameter of the model, it is not identifiable. A common assumption employed by users of MNL models is that the scale parameter, μ , is equal to one, which has a homoscedastic disturbance (Garber-Yonts, 2001). Empirical observations about this assumption found that it was not significantly different than one (Xu, 1997; Adamowicz et al., 1998). Therefore, we adhere to this assumption in this study. The log likelihood function for the MNL model can be written as Equation 6:

$$\ln L = \sum_n \sum_{i \in A_n} s_{in} P(i/A_n) = \sum_n \sum_{i \in A_n} s_{in} [\beta' f(X_{in}, H_n) - \ln \sum_{j \in A_n} \beta' f(X_{jn}, H_n)] \quad (6)$$

Where $s_{in} = 1$ if alternative i is chosen by individual n , otherwise $s_{in} = 0$. Garber-Yonts (2001) provides the details explanation about the derivatives of all Equations related to MNL. The necessary first order conditions to maximize the likelihood function are obtained by setting the first derivative of Equation 6 with respect to the parameter vector equal to zero:

$$\frac{\partial}{\partial \beta} \ln L = \sum_n \sum_{i \in A_n} s_{in} \left[f(X_{in}, H_n) + \frac{\sum_{j \in A_n} e^{\mu \beta' f(X_{in}, H_n)} (f(X_{in}, H_n))}{\sum_{j \in A_n} e^{\mu \beta' f(X_{in}, H_n)}} \right] = 0 \quad (7)$$

Estimation of the parameters of this model can be done by using maximization of the multinomial likelihood. This usually requires numerical procedures, and Fisher scoring or Newton-Raphson often work rather well. McFadden (1974) argues that, under certain conditions, $\ln L$ in Equation 6 is globally concave so that a solution to Equation 7 exists and is unique. Thus the maximum likelihood estimator of β is consistent, asymptotically normal, and asymptotically efficient.

Estimation of Hicksian welfare effects from the MNL choice probabilities follows the method outlined by Hanemann (1984) and Hanemann and Kanninen (1999). Given a quantity change in the level of a public good from X^0 to X^1 , the compensating surplus which exactly offsets the utility gain of the change is the level of B which provides the equality:

$$v(p, X^1, y - B, H, \varepsilon) = v(p, X^0, y, H, \varepsilon) \quad (8)$$

where v is indirect utility, p is the vector of market prices, a X is vector of attributes other than the bid level B , y is income, H is a vector of the socio-demographic characteristics, and ε is a random error term. The objective is to obtain the solution for the expected value of $B = B(p, X^0, X^1, y, H, \varepsilon)$, which is the maximum WTP for the change from X^0 to X^1 . Assuming the additive separability of the cost attribute of the individual's indirect utility function, we can express the deterministic part of utility as shown in Equation 9:

$$v_{in} = \beta' f(X_{in}, H_n) + \alpha B_{in} \quad (9)$$

Where, B is the specified bid level alternative i , and α is associate parameter. The following measures Total WTP/Total WTA (TWTP/TWTA) for a change in the attributes of a good from state i to state j aggregated over all observations (Hanemann, 1984; Adamowicz et al., 1994; Xu, 1997; Garber-Yonts, 2001):

$$\begin{aligned}
 TWTP/TWTA = CS &= -\frac{1}{\alpha} \left\{ \ln \sum_{i \in A} e^{v_{in}} - \ln \sum_{j \in A} e^{v_{jn}} \right\} \\
 &= -\frac{1}{\alpha} \left\{ \ln \sum_{i \in A} e^{\beta' f(X_{in}, H_n) + \alpha B_{in}} \right\} - \left\{ \ln \sum_{j \in A} e^{\beta' f(X_{jn}, H_n) + \alpha B_{jn}} \right\}
 \end{aligned} \tag{10}$$

If the mean value of TWTP/TWTA for the change in all attributes from state i to state j is for interest, Equation 10 simplifies to:

$$TWTP/TWTA = CS = -\frac{1}{\alpha} [\alpha(B_i - B_j) + \beta' (f(X_i, \bar{H}) - f(X_j, \bar{H}))] \tag{11}$$

where $f(X, H)$ is evaluated at the sample mean value of H , recalling that H drops out of the Equation if it enters $f(\cdot)$ additively. The TWTP/TWTA for the “part-worth” of the change of an individual attribute k from state i to state j , holding other attributes constant, further simplifies to Equation 12:

$$TWTP/TWTA_k = CS_k = -\frac{\beta_k}{\alpha} [f(X_{ik}, \bar{H}) - f(X_{jk}, \bar{H})] \tag{12}$$

Finally, as adopted by Hanemann et al. (1991); Xu (1997) and Garber-Yonts, (2001) the Hicksian compensated demand curve, depicting marginal WTP/WTA for attribute k at level i , is given as Equation 13:

$$TWTP/TWTA_k = \frac{\partial}{\partial X} CS = -\frac{\beta}{\alpha} f(X_{ik}, \bar{H}) \tag{13}$$

In choice modelling applications, different components of specific public good as well as monetary factors should be included as attributes of the options in a choice set. Thus, choice modelling allows one to obtain compensating surplus estimates so that one can account for the welfare change generated by a bundle of changes in relevant attributes. It is also possible to determine the relative importance of these attributes to people in making their choices. Haneman and Kanninen (1999) make an important distinction between the conventional regression techniques used in analysis of open ended WTP data and the limited dependent variable models used in conjunction with discrete

choice elicitation methods. With the former, the investigator obtains an estimate of the mean WTP conditional on the regressors. The later estimates the entire conditional cumulative distribution function (cdf) of the dependent variable. The preferred measure of central tendency by which to summarize the estimated cdf is therefore at the discretion of the investigator, and its selection can significantly alter the results of the analysis (Garber-Yonts, 2001).

It is clear now that the choice experiment technique is an application of the characteristics theory of value combined with random utility theory (see, for example, Thurstone, 1927; Lancaster, 1966; Manski, 1977). In this method, respondents are asked to choose between different bundles of (environmental) goods, which are described in terms of their attributes, or characteristics, and the levels that these take. The CE approach is essentially a structured method of data generation. It relies on carefully designed choice tasks that help reveal the factors influencing choice. Designing a CE technique requires careful definition of the attribute levels and ranges. Furthermore, the choice experiment approach involves the use of statistical design theory to construct choice scenarios which can yield parameter estimates that are not confounded by other factors. In the next section, we discuss the main steps that we should follow when applying CE method for environment valuation.

CHOICE EXPERIMENT (CE) METHOD

As mentioned in the previous section, the CE method has its theoretical grounding in Lancaster's model of consumer choice (Lancaster, 1966). Lancaster proposed that consumers derive satisfaction not from goods themselves, but from the attributes they provide. To illustrate the basic model behind choice experiments, assume that particular household has a utility function of the form:

$$U_{ij} = U(X_{ij}, Z_i) \quad (14)$$

Where, for any household i , a given level of utility will be associated with any alternative of the commodity j . Utility derived from any alternatives depend on the attributes of the commodity X_{ij} and the social and economic characteristics of the household Z_i , since different household may receive different levels of utility from these attributes.

According to the random utility model, the utility of a choice comprises of a systematic (deterministic) component, T_{ij} and an error (random) component, e_{ij} , which is independent of the deterministic part and follows a predetermined distribution (Hanemann et al., 1991):

$$U_{ij} = T_{ij} + e_{ij} \quad (15)$$

The systematic component can be explained as a function of the characteristics of the commodity and of the social and economic characteristics of the household. Accordingly, Equation 15 can be expressed as $U_{ij} = T(X_i, Z_i) + e_{ij}$. Given an error part in the utility function, predictions cannot be made with certainty and the analysis becomes one of probabilistic choice (Bateman et al., 2003). Consequently, choices made between alternative commodities will be a function of the probability that the utility associated with a particular commodity option (j) is higher than that for other alternative commodities. Hence, the probability that household i will choose commodity j over all other options n is given by:

$$P_{ij} = \text{prob}\{T_{ij} + e_{ij} > T_{in} + e_{in}\} \quad \text{Where, } j \neq n.$$

We assume that the relationship between utility and attributes follows a linear path in the parameters and variables. We further assume that the error terms are identically and independently distributed with a Weibull distribution² (Greene, 1997). These assumptions ensure that the probability of any particular alternative j being chosen can be expressed in terms of logistic distribution. This specification is known as the CL model (McFadden, 1974; Greene, 1997; Maddala, 1999) which has the following general form:

$$P_{ij} = \frac{\exp(X_i' \beta + Z_i' \alpha)}{\sum_{j=1}^J \exp(X_i' \beta + Z_i' \alpha)} \quad (16)$$

² Weibull distribution is a continuous probability distribution. For further details about the basic properties of this distribution, please see Greene (1997).

The components of X_{ij} are typically called the attribute of the choices. However, Z_i contains characteristics of the individual and is, therefore, the same for all choices. Equation 16 is the probabilistic response function and it shows that, given all other options the probability that household i selecting the option j type commodity. The CL model generates results for a conditional indirect utility function of the form:

$$T_{ij} = \beta + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \alpha_1 Z_1 + \alpha_2 Z_2 + \dots + \alpha_k Z_k \quad (17)$$

where β is the alternative specific constant (ASC), that captures the effects in utility from any attributes not included in choice specific attributes (Rolfe et al., 2000). The number of attributes of the commodity considered is m and the number of social and economic characteristics of the household to explain the choice of the commodity is k . The vectors of coefficients are attached to the vector of attributes (X) and to a vector of socio-economic factors (Z) that influence utility, respectively.

The CE method is consistent with utility maximization and demand theory (Bateman et al., 2003). When parameter estimates are obtained, welfare measures can be estimated from the CL model using the following formula:

$$CS = \frac{\ln \sum_i \exp(T_{i1}) - \ln \sum_i \exp(T_{i0})}{\delta} \quad (18)$$

where CS is the compensating surplus welfare measure, δ is the marginal utility of income (generally represented by the coefficient of the monetary attribute in the CE) and T_{i0} and T_{i1} represent indirect utility functions of alternative i (with subscript 0 indicating the base situation and 1 indicate the changed situation) before and after the change under consideration. For the linear utility index, the marginal value of change within a single attribute can be represented as a ratio of coefficients, reducing Equation 18 to 19:

$$W = - \left(\frac{\beta_{attribute}}{\beta_{monetary\ attribute}} \right) \quad (19)$$

Equation 19, the implicit prices (W) for the various attributes can be calculated. These demonstrate the marginal rate of substitution between

cost and the attribute in question. This is the same as the marginal welfare measure (WTP or WTA) for a change in any of the attributes.

An alternative model specification to the CL model is RPL model which is increasingly becoming popular in CE studies. The advantage of RPL model is that it accounts for consumers' taste heterogeneities and also relaxes the Independence of Irrelevant Alternatives (IIA) assumption of the CL model. It also provides a flexible and computationally practical econometric method for any discrete choice model derived from random utility maximisation (McFadden and Train, 2000). More importantly preferences are in fact heterogeneous and accounting for this heterogeneity enables estimation of unbiased estimates of individual preferences and enhances the accuracy and reliability of estimates of parameters of the model and total welfare (Greene, 1997). Furthermore, accounting for heterogeneity enables prescription of policies that take equity concerns into account. This is because an understanding of who will be affected by a policy change in addition to understanding the aggregate economic value associated with such changes is necessary (Boxall and Adamowicz, 2002). Formally, the random utility function in the RPL model is given by:

$$U_{ij} = U[X_j(\beta + \mu_i), Z_i] \quad (20)$$

Similarly, to the CL model, indirect utility is assumed to be a function of the choice attributes (X_j), with parameters β , which due to preference heterogeneity may vary across respondents by a random component μ , and of the social, economic and attitudinal characteristics (Z_i), namely income, education, household size and attitudes towards the relevant good or service. By accounting for unobserved heterogeneity, Equation 16 now becomes:

$$P_{ij} = \frac{\exp[X_{ij}'(\beta + \mu_i) + Z_i'\alpha]}{\sum_{j=1}^J \exp[X_{ij}'(\beta + \mu_i) + Z_i'\alpha]} \quad (21)$$

Since this model is not restricted by the IIA assumption, the stochastic part of utility may be correlated among alternatives and across the sequence of choices via the common influence of μ_i . Treating preference parameters as random variables requires estimation by simulated maximum likelihood (Kikulwe et al., 2011). In general, the maximum likelihood algorithm searches for a solution by simulating n draws from distributions with given means and standard deviations. Probabilities are calculated by integrating the joint simulated

distribution. Recent applications of the RPL model have shown that this model is superior to the CL model in terms of overall fit and welfare estimates (Breffle and Morey, 2000; Layton and Brown, 2000; Carlsson et al., 2003; Kontoleon, 2003; Lusk et al., 2003; Morey and Rossmann, 2003).

Even if unobserved heterogeneity can be accounted for in the RPL model, the model fails to explain the sources of heterogeneity (Boxall and Adamowicz, 2002). This can be done by including interactions of respondent-specific social, economic and attitudinal characteristics with choice specific attributes and/or with ASC in the utility function. This enables the RPL model to pick up preference variation in terms of both unconditional taste heterogeneity (random heterogeneity) and individual characteristics (conditional heterogeneity), and hence improve model fit (e.g. Revelt and Train, 1998; Morey and Rossmann, 2003; Kontoleon, 2003). In the context of empirical application of choice experiment model, choice experiment design as well as model selection steps are extremely important. Therefore, the next section discusses basic steps of choice experiment design and selecting the appropriate model for econometric estimation.

CHOICE EXPERIMENT DESIGN AND MODEL SELECTION

In the CE method³, respondents are presented with panels of choices with two or more alternatives each, where each alternative is a bundle of attributes which are specified at different levels in each alternative (Louviere et al., 2000). The inclusion of a price or cost attributes permits estimating the effect of cost on the respondents' choice. For example, if we consider farmers' preference for different type of farms, a farmer may choose from a number of different farms in her choice set, each of which exhibits variation in an array of attributes such as crops diversity, livestock diversity, mix farming system, landrace cultivation and organic production. A farmer chooses the type of farm in a given season depending on the balance of preferences for different attributes and the degree to which they are represented at a given farm. In a survey context, the researcher should identify the essential attributes and levels of the environmental goods in question and designs the

³ For a detailed explanation of choice experiment design techniques, please see Louviere et al. (2000), Bennet and Blamey (2001) and Bateman et al. (2002)

choice question to reveal the structure of the respondents' preferences (Bateman et al., 2002).

Adamowicz et al. (1999) provided several stages of designing a CE study. They are as follows:

1. Identification of relevant attributes
2. Selection of measurement unit for each attribute
3. Specification of the number and magnitude of the attribute levels
4. Experimental design
5. Model estimation
6. Use of parameters to simulate choice

The first three steps are involved in developing a concise and sufficiently complete representation of the valuation scenario which will provide the survey respondent with appropriate information set on which to base statements of preference. This phase uses information obtained from secondary sources, experts in the field, focus groups and personal interviews in order to refine the informational content of the survey instrument. The selection of attributes in relation to the choices of interest is very important in framing a CE exercise. According to Blamey et al. (2000) attribute selection needs to take place from both the perspectives of the end-user (the population of interest) and the decision-makers/resource managers to ensure that the attributes are not only easily identifiable, but produce policy-relevant information. Another goal of the attribute selection process is to minimize the number of attributes as the use of a large number of attributes is likely to lead to lower data reliability due to the excessive cognitive burden it would place on respondents (Mogas et al., 2002). Identification of appropriate attribute ranges is another basic framing task in choice experiment, as a failure to accept trade-offs indicates that the range of attribute levels offered is not salient (Johnson et al., 2000). In determining how many attributes to include in a study design, there is often a trade-off between describing trade-offs accurately (requiring more attributes) and minimizing choice and experimental design complexity (requiring fewer attributes). Louviere and et al. (1993) claims to have successfully administered surveys with up to 32 choice tasks though this requires scaling down the number of alternatives and attribute levels accordingly. Boxall et al. (2002) suggests that respondents can endure large numbers of choice sets but sets with more than six alternatives tend to exceed cognitive limits. Louviere et al. (1993) suggest that the average choice experiment survey employs

seven attributes, four choice sets and four alternatives per set, though they note that there is a great deal of variability and this average does not constitute a best practice.

After identifying the attributes for a particular experiment, the analyst must assign values or levels to each attribute. These levels should be chosen to represent the relevant range of variation in the present or future interest of respondents. In general, focus group discussions will indicate the level of the attributes as well as the best way to present them. Though commonly presented in words and numbers, attribute levels may be presented using pictures. To the extent that visual representations of attribute levels are utilised, it is likely that respondents will perceive levels more homogeneously, likely leading to more precise parameter estimates in the modelling stage (Alpizar et al., 2001).

Experimental design⁴ is the next important aspect of choice modelling and it is concerned with how to create the choice sets in an efficient way or how to combine attribute levels into profiles of alternatives and profiles into choice sets. In practice, a design is developed in two steps: (i) obtaining the optimal combinations of attributes and attribute levels to be included in the experiment and (ii) combining those profiles into choice sets. A starting point is a full factorial design, which is a design that contains all possible combinations of the attribute levels that characterize the different alternatives. A full factorial design is, in general, very large and not tractable in a choice experiment (Louviere et al., 2000). Therefore, we need to choose a subset of all possible combinations, while following some criteria for optimality and then construct the choice sets. The standard approach used in most research has been to use orthogonal designs, where the variations of the attributes of the alternatives are uncorrelated in all choice sets. More recently researchers in marketing have developed design techniques based on the Doptimal criteria for non-linear models in a choice experiment context. However, there can be some problems with these more advanced design strategies due to their complexity, and it is not clear whether the advantages of being more statistically efficient outweigh the problems (Scarpa and Rose, 2008)⁵.

⁴ This step is much more complex in choice experiment in that the experimental design is critical to producing a data set that will yield estimable parameters for the attributes in an econometric model of preferences.

⁵ For example, utility balance in more advanced design makes the choice harder for the respondents, since they have to choose from alternatives that are very close in terms of utility.

As mentioned above, the most well-known fractional factorial design type is the orthogonal design, which aims to minimise the correlation between the attribute levels in the choice situations (Kuhfeld, 2005). However, these orthogonal designs have limitations and cannot avoid choice situations in which a certain alternative is clearly more preferred over the others (hence not providing much information). More recently, several researchers have suggested another type of fractional factorial designs, so-called efficient designs (Hensher et al., 2005; Scarpa and Rose, 2008). Instead of merely looking at the correlation between the attribute levels, efficient designs aim to find designs that are statistically as efficient as possible in terms of predicted standard errors of the parameter estimates. Essentially, these designs attempt to maximise the information from each choice situation. In case any information about the parameters is available, then efficient designs will always outperform orthogonal designs (Kessels et al., 2006). This is due to the fact that efficient designs use the knowledge of the prior parameters to optimise the design in which the most information is gained from each choice situation (e.g. dominant alternatives can be avoided as the utilities can be computed). While efficient designs outperform the orthogonal designs, prior parameter estimates need to be available (Hensher et al., 2005). Therefore, efficient designs rely on the accuracy of the prior parameter estimates.

Three reasons can be given to justify using orthogonal design in a particular study. Firstly, it allows for an independent estimation of the influence of each design attribute on choice. Secondly, with the absence of prior parameter, there is no way to apply efficient design in the study. Thirdly, the common use of orthogonal designs in stated choice studies is largely a result of historical impetus. In the past, the experimental design literature has been primarily concerned with linear models (such as linear regression models), where the orthogonality of data is considered important (Scarpa and Rose, 2008). In linear regression models, this is because (a) orthogonality ensures that the model will not suffer from multicollinearity, and (b) orthogonality is thought to minimise the variances of the parameter estimates, which are taken from the variance-covariance (*VC*) matrix of the model (Hensher et al., 2005). The *VC* matrix of a linear regression model is given in Equation 22.

$$VC = \sigma^2 [X'X]^{-1} \quad (22)$$

where σ^2 is the model variance, and X is the matrix of attribute levels in the design or in the data use to estimate. Fixing the model variance,

the elements of the VC matrix for linear regression models are minimised when the X matrix is orthogonal. A design that results in a model where the elements contained within the VC matrix are minimised is preferable, for two reasons (Hensher et al., 2005). Firstly, such a design will produce the smallest possible standard errors, and hence maximise the t -ratios produced from that model. Secondly, an orthogonal design will produce zero-off diagonals in the models VC matrix, thus ensuring that the parameter estimates are unconfounded with one another (or no multicollinearity problem). As such, orthogonal designs, at least in relation to linear models, meet the two criteria for a good design (Scarpa and Rose, 2008). They allow for an independent determination of each attributes contribution on the dependent variable, and they maximise the power of the design to detect statistically significant relationships (e.g. maximise the t -ratios at any given sample size).

The next step of choice experiment involves econometric model selection and estimation. The most common model estimated in economics literature has been the MNL model, and the most common estimation criterion is maximum likelihood. The MNL model is easy to estimate, and interpretation is straightforward. However, there are also examples of other choice model specifications such as CL model and RPL model. Selection between MNL and CL depends on whether the researcher is interested in including socioeconomic variables in addition to the choice attribute into the model. If researcher uses only choice attributes, MNL model can give higher accuracy of the model fits. However, if the researcher uses choice attributes as well as socioeconomic variables into the model, CL model provides more accurate results (Rolfe et al., 2000). In empirical setting, inclusion of social and economic characteristics is also beneficial in avoiding IIA violations, since social and economic characteristics relevant to preferences of the respondents can increase the systematic component of utility while decreasing the random error (Rolfe et al., 2000; Bateman et al., 2003).

The MNL model relies on the assumption of the independence of irrelevant alternatives⁶. The IIA arises from the assumption about the IID of the error term. IID of error term means that it has an extreme value error distribution. The IIA means that the probability of choosing

⁶ The independence of irrelevant alternatives means that, all else being equal, a person's choice between two alternative outcomes is unaffected by what other choices are available.

an alternative is dependent only on the options from which a choice is made, and not on any other options that may exist. If the IIA/IID is violated, the estimates derived from the model could be biased and not generate accurate values for inclusion in cost benefit analysis (Ben-Akiva and Lerman, 1985). The IIA property allows the addition or removal of an alternative from the choice set without affecting the structure or parameters of the model. This assumption has three main advantages. Firstly, the model can be estimated and applied in cases where different members of the population face different sets of alternatives. Secondly, this property simplifies the estimation of the parameters in the MNL and CL models. Third, this property is advantageous when applying a model to the prediction of choice probabilities for a new alternative. On the other hand, the IIA property may not properly reflect the behavioural relationships among groups of alternatives (Hensher et al., 2005). That is, other alternatives may not be irrelevant to the ratio of probabilities between a pair of alternatives. In some cases, this will result in erroneous predictions of choice probabilities.

There are various reasons why IIA/IID violation could occur. One possibility is the existence of random taste variations (that is heterogeneity). To account for this, a model which includes socioeconomic variables in addition to the attributes in the choice sets can be estimated (Bennett and Blamey, 2001). The socio-economic information could be included in two different ways. The first is by interactions with the attributes in the choice sets. The second method includes the socio-economic information through interactions with the alternative specific constants. These interactions show the effect of various socio-economic characteristics on the probability that a respondent will choose particular options.

Alternative model specifications to MNL models are CL and RPL. The CL model allows us to estimate the effect of choice-specific variables on the probability of choosing a particular alternative. The CL model also assumes the IIA property, which states that the relative probabilities of two options being chosen are unaffected by introduction or removal of other alternatives. In other words, the probability of a particular alternative being chosen is independent of other alternatives. If the IIA property is violated then CL model results will be biased and hence a discrete choice model that does not require the IIA property, such as the RPL model, should be used. To test whether the CL model is appropriate, the Hausman and McFadden (1984) test for the IIA property can be employed. In this case, whether or not IIA property

holds can be tested by dropping an alternative from the choice set and comparing parameter vectors for significant differences. A RPL model is a generalization of a standard multinomial logit. The advantages of a RPL model are that (i) the alternatives are not independent (the model does not exhibit the independence of irrelevant alternatives property) and (ii) there is an explicit account for unobserved heterogeneity.

CHOICE EXPERIMENT SURVEY

Under the CE method a sample of people is asked to choose their most preferred alternatives from a sequence of grouped options that relate to different management strategies. Each option is described in terms of its outcomes and a personal monetary cost to be borne personally by the respondent. In general, survey is the common technique that is used to collect data. The questionnaire is usually a paper and pencil task that is presented through an interviewer. While its main content will be different choice scenarios through which the respondent will be guided, it may also include sections requesting socio-demographic, economics, and attitudinal and past behaviour data⁷.

In general, the questionnaire needs to be developed using the results from focus groups' discussions and a pre-test. The purposes of the focus group studies are to determine attributes relevant to respondents and policy makers and test a draft questionnaire. Also before the interview starts it is required to confirm whether the respondents are generally those responsible for decision making. In face-to-face interviews, each respondent can be presented with several choice sets showing various options. Before answering the choice sets, respondents need to be requested to keep in mind their available income, food consumption expenditure, available labour, size of the land and other things on which they may consider when making a decision. It is obvious that the CE part is the most important section of the questionnaire and it needs expert knowledge and careful attention. In a CE, individuals are presented with a choice set or series of choice sets that are framed with various attributes and attribute levels and are asked to choose one bundle at a varied set of price and attribute levels. Consumers' willingness to accept (WTA) compensation payment for each attribute is then computed from estimates of econometric models. An intrinsic

⁷ Socio economic aspects such as community, gender, age, marital status, literacy level, income, expenditure, savings and indebtedness provide a base for studying the impact of any program.

problem that all researchers face in designing a survey questionnaire is how much information or complexity to incorporate. Specifically, these issues may include which attributes should be used, how many levels of each attribute need to be considered, how many alternatives need to be presented in each choice set, and how many choice sets should be included in each questionnaire. The process for designing CE questions must be able to answer all these questions carefully.

CONCLUSION

The overall objective of this paper is to explain the basic steps of the CE method. CE study estimates the possible benefits that could be achieved from changing existing scenario to a new scenario. Under this method a sample of people is asked to choose their most preferred alternatives from a sequence of grouped options that relate to different management strategies. Each option is described in terms of its different outcomes and a monetary cost to be borne personally by the respondent. By analysing the choices made by respondents it is possible to infer the trade-offs that people are willing to make between money and greater benefits of changing the existing situation. A choice experiment is a highly structured method of data generation, relying on carefully designed tasks (experiment) to reveal the factors that influence choices. Experimental design theory is used to construct profiles of the environmental good in terms of its attributes and levels of these attributes. Profiles are assembled in choice sets, which are in turn presented to the respondents, who are asked to state their preferences. In a well-designed CE study, we need to follow all these steps explained in this article in order to increase the accuracy as well as reliability of the results of the study. We need to carefully design the choice experiment survey and used appropriate econometric techniques for the analysis.

REFERENCES

- Adamowicz, W., J. Louviere and M. Williams. 1994. Combining stated and revealed preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management*, 26(3): 271-292.
- Adamowicz, W., Boxall, P. C., J. Louviere and M. Williams. 1998. Stated preference approaches for measuring passive use values: choice experiments and contingent valuation. *American Journal of Agricultural Economics*, 80(1): 65-75.
- Adamowicz, W. L., Boxall, P. C., Louviere, J., J. Swait and M. Williams. 1999. Stated-preference methods for valuing environmental amenities. In: *Valuing environmental preferences: Theory and practice of the contingent valuation method in the US, EU, and developing countries*, I. J. Bateman and K. G. Willis (Eds.), Oxford University Press, pp: 460-479.
- Bateman, I. J., Carson, R. T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Ozdemiroglu, E., Pearce, D. W., R. Sugden and J. Swanson. 2002. *Economic Valuation with Stated Preference Techniques*. Cheltenham: Edward Elgar.
- Bateman, I. J., Carson, R. T., Day, B. W., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Ozdemiroglu, E., Pearce, D. W., R. Sugden and S. Swanson. 2003. *Guidelines for the Use of Stated Preference Techniques for the Valuation of Preferences for Non-Market Goods*. Cheltenham: Edward Elgar.
- Ben-Akiva, M. and S. R. Lerman. 1985. *Discrete Choice Analysis. Theory and Application to Travel Demand*. The MIT Press, Cambridge.
- Bennett, J. and R. Blamey. 2001. *The Choice Modelling Approach to Environmental Valuation*, Edward Elgar, Cheltenham, UK.
- Bishop, R. and D. Romano. 1998. *Environmental Resource Valuation Applications of the Contingent Valuation Method in Italy*. Kluwer Academic Publishers, Boston.

- Blamey, R., Rolfe, J., J. Bennett and M. Morrison. 2000. Valuing remnant vegetation in Central Queensland using choice modelling. *The Australian Journal of Agricultural and Resource Economics*, 44(3): 439-456.
- Boxall, P. C. and W. L. Adamowicz. 2002. Understanding heterogeneous preferences in random utility models: A latent class approach. *Environmental and Resource Economics*, 23: 421- 446.
- Breffle, W. and E. R. Morey. 2000. Investigating preference heterogeneity in a repeated discrete choice recreation demand model of Atlantic salmon fishing. *Marine Resource Economics*, 15(1): 1-20.
- Cameron, A. C. and P. K. Trivedi. 1998. Regression Analysis of Count Data, Econometric Society Monograph No.30: Cambridge University Press.
- Carlsson, F., P. Frykblom and C. Liljenstolpe. 2003. Valuing wetland attributes: An application of choice experiments. *Ecological Economics*, 47(1): 95-103.
- Champ, P. A., T. C. Brown and K. J. Boyle. 2004. A Primer on Nonmarket Valuation. Dordrecht: Kluwer Academic Publishers.
- Freeman III, A. M. 2003. The Measurement of Environmental and Resource Values: Theory and Methods, (2nd Eds.), Resources for the Future, Washington, D.C.
- Garber-Yonts, B. 2001. A Choice Experiment Analysis of Public Preferences for Conservation of Biological Diversity in the Oregon Coast Range, Unpublished Doctoral Dissertation, Oregon State University, USA.
- Greene, W. H. 1997. Econometric Analysis. (3rd Eds.), New Jersey: Prentice Hall.
- Hanemann, M. 1984. Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics*, 66(3): 332-341.

- Hanemann, M. and B. Kanninen. 1999. The statistical analysis of discrete-response CV data. In: Valuing Environmental Preferences, Bateman, I. and K. Willies (Eds.), Oxford University Press.
- Hanemann, W. M., J. Loomis and B. Kanninen. 1991. Statistical efficiency of double bounded dichotomous choice contingent valuation. *American Journal of Agricultural Economics*, 73: 1255-1263.
- Hanley, N., MacMillan, D., Wright, R. E., Bullock, C., Simpson, I., D. Parrisson and B. Crabtree. 1998. Contingent valuation versus choice experiments: estimating the benefits of environmentally sensitive areas in Scotland. *Journal of Agricultural Economics*, 49(1): 1-15.
- Hausman, J. and D. McFadden. 1984. Specification tests for the multinomial logit Model. *Econometrica*, 52(5): 1219-1240.
- Hensher, D. A., J. M. Rose and W. H. Greene. 2005. Applied Choice Analysis: A Primer. Cambridge University Press, Cambridge, UK.
- Johnson, F. R., K. E. Mathews and M. F. Bingham. 2000. Evaluating welfare-theoretic consistency in multiple-response, stated-preference surveys. Triangle Economic Research Technical Working Paper, No T-0003. Triangle Economic Research, Durham.
- Kikulwe, E. M., Birol, E., J. Wesseler and J. Falck-Zepeda. 2011. A latent class approach to investigating demand for genetically modified banana in Uganda. *Agricultural Economics*, 42 (95): 547-560.
- Kontoleon, A. 2003. Essays on Non-Market Valuation of Environmental Resources: Policy and Technical Explorations. Unpublished Doctoral Dissertation, University of London.
- Lancaster, K. 1966. A new approach to consumer theory. *Journal of Political Economy*, 74(2): 132-157.
- Layton, D. F. and G. Brown. 2000. Heterogeneous preferences regarding global climate change. *The Review of Economics and Statistics*, 82(4): 616-624.

- Li-zhi, G. 2003. The conservation of Chinese rice biodiversity: genetic erosion, ethnobotany and prospects. *Genetic Resources and Crop Evolution*, 50(1): 17-32.
- Louviere, J., M. Fox and W. Moore. 1993. Cross-task validity comparisons of stated preference choice models. *Marketing Letters*, 4: 205-213.
- Louviere, J. J., Hensher, D. A., J. D. Swait and W. Adamowicz. 2000. *Stated Choice Methods: Analysis and Applications*. Cambridge University Press, Cambridge.
- Luce, D. 1959. *Individual Choice Behavior*. New York: John Wiley.
- Lusk, J. L., J. Roosen and J. A. Fox. 2003. Demand for beef from cattle administered growth hormones or fed genetically modified corn: a comparison of consumers in France, Germany, the United Kingdom and the United States. *American Journal of Agricultural Economic*, 85(1): 16-29.
- Maddala, G. S. 1999. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge.
- Manski, C. 1977. The structure of random utility models. *Theory and Decision*, 8(3): 229-254.
- McFadden, D. 1974. Conditional Logit Analysis of Qualitative Choice Behaviour. In: *Frontiers in Econometrics*, Zarembka, P. (Eds.), New York: Academic Press.
- McFadden, D. and K. Train. 2000. Mixed MNL models of discrete response. *Journal of Applied Economics*, 15(5): 447-470.
- Mogas, J., P. Riera and J. Bennett. 2002. A comparison of contingent valuation and choice modeling: estimating the environmental values of Catalonian forests. Occasional paper no. 1, National Centre for Development Studies, Australian National University.

- Morey, E. and K. Rossmann. 2003. Using stated-preference questions to investigate variations in willingness to pay for preserving marble monuments: classic heterogeneity, random parameters, and mixture models. *Journal of Cultural Economics*, 27(3-4): 215-229.
- Revelt, D. and K. Train. 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. *The Review of Economics and Statistics*, 80(4): 647-657.
- Rolfe, J., J. J. Bennett and J. Louviere. 2000. Choice modelling and its potential application to tropical rainforest preservation. *Ecological Economics*, 35(2): 289-302.
- Scarpa, R. and J. M. Rose. 2008. Designs efficiency for nonmarket valuation with choice modelling: how to measure it, what to report and why. *The Australian Journal of Agricultural and Resource Economics*, 52(3): 253-282.
- Thurstone, L. L. 1927. A law of comparative judgment. *Psychological Review*, 34 (4): 278-286.
- Xu, Y. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25(1): 61-83.