Discussion Papers In Economics And Business

A Review of Stated Choice Method

Junyi Shen

Discussion Paper 05-27

Graduate School of Economics and
Osaka School of International Public Policy (OSIPP)
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

A Review of Stated Choice Method

Junyi Shen*

October 2005

Abstract

This paper reviews Stated Choice Method (SCM), paying particular attentions on its theoretical background, application, empirical models, experimental design, and procedure to execute. The review suggests that comparing to other stated preference (SP) methods, SCM has a major advantage that it meets the objective of a stated preference analysis to simulate actual consumer behavior by allowing simultaneous evaluations of a number of alternatives or a choice between alternatives. Some advanced models based on the degrees of relaxation of the Independently and Identically Distributed (IID) assumption on error terms are introduced. More complex model seems to be more plausible than relatively simple specifications. Two tests for nested and non-nested models are also discussed to help judge that one model is superior to another model. Finally, this paper introduces the procedure of executing a Stated Choice (SC) experiment.

Keywords: Stated Choice Method (SCM), Stated Preference (SP) method, Independent and Identical Distribution (IID), Extreme Value type I (EV1) distribution

JEL classifications: C35, C81, C93

---

*The author would like to thank Yoshizo Hashimoto for his helpful comments. Financial support from the Japanese Ministry of Education, Culture, Sports, Science and Technology (Grant-in-Aid for Scientific Research A(2) 16203020) is gratefully acknowledged. All views expressed here and any errors are the sole responsibility of the author.

Correspondence to: Osaka School of International Public policy, Osaka University,
1-31 Machikaneyama, Toyonaka, Osaka 560-0043, Japan. Tel/Fax: +81-6-68505652
E-mail: junyi@osipp.osaka-u.ac.jp

1. Introduction

Stated Choice Method (SCM) is a research technique in the family of Stated Preference (SP) methods. In stated preference studies, information about decision makers' preferences is elicited by using specifically designed hypothetical situation. Hence the data generated in stated preference studies are derived from decision experiments, which is the main difference from an analysis of revealed preferences. In Reveal Preference (RP) studies, such as those using the travel cost method and the hedonic pricing method, decision makers' preferences are revealed in their decisions in real choice situations. There are various reasons why a stated preference study may be preferred to an analysis of preferences that are revealed in actual choices. Louviere *et al.* (2000, p.21-22), for example, mentioned that:

> "Despite well-developed economic theory for dealing with real market choices, there are a number of compelling reasons why economists and other social scientists should be interested in *stated preference* (SP) data, which involve choice responses from the same economic agents, but evoked in hypothetical (or virtual) markets:
>
> - Organisations need to estimate demand for new products with new attributes or features.
> - Explanatory variables have little variability in the marketplace.
> - Explanatory variables are highly collinear in the marketplace.
> - New variables are introduced that now explain choices
> - Observational data cannot satisfy model assumptions and/or contain statistical 'nasties' which lurk in real data.
> - Observational data are time consuming and expensive to collect.
> - The product is not traded in the real market."

Similarly, Kroes and Shelden (1988, p.13) also stated that:

> "It [reveal preference method] is against the backdrop of such problems that the use of stated preference methods became an attractive option in transport research. Broadly, these methods are easier to control (because the researcher defines the conditions which are being evaluated by the respondents); they are more flexible (being capable of dealing with a wider variety of variables); and they are cheaper to apply (as each respondent provides multiple observations for variations in the explanatory variables which interest the analyst)."

A stated choice survey employs a carefully designed questionnaire in which respondents are given a sequence of questions or choice sets. In each choice set, they are asked to indicate their preferred option from a set of alternatives. Each alternative option is described in terms of a number of key attributes that are specified at different levels. The configuration of attribute levels that describe the alternatives follows an experimental design and varies between choice sets. The response data, which usually also include individuals' socio-economic characteristics, enable not only the estimation of the relationships between attribute levels and the choice probabilities, but also the estimation of the extent of the trade-offs between the attributes made by respondents.

The SCM is one of a number of different stated preference methods. Others include the contingent valuation method (CVM) and conjoint rating and ranking. A major advantage of SCM compared with the other stated preference methods is that it meets the objective of a stated preference analysis to simulate actual consumer behavior. It pertains to respondents making a choice between a number of alternatives on offer. This is in contrast to CVM that has been applied to derive welfare estimates in the context of non-market environmental values. In CVM, respondents are asked to evaluate a "current situation" and one alternative option only, and to indicate their willingness to pay (WTP) for the change in the environmental situation. However, because it does not elicit choices, and because it does not involve the simultaneous evaluation of various options, CVM is not an appropriate method for choice-based analyses. Another concern with comparison between CVM and SCM is that CVM relies very heavily on the accuracy of descriptions, in contrast, SCM relies less on the accuracy and completeness of any particular alternative, but more on the accuracy and completeness of the product characteristics and attributes used to describe alternatives (Louviere *et al.* 2000, ch.12).

Conjoint analysis is the generic term for the attribute-based analysis of consumer decision making (see Green and Srinivasan 1978). Respondents are asked to evaluate various options that are described in terms of a set of attributes[1]. Conjoint methods include conjoint rating and conjoint ranking[2]. In a conjoint rating study, respondents are asked to rate their likelihood of purchasing a particular attribute combination. Using multiple ratings data for each respondent, the relationship between the ratings and the individual attributes can be established in a regression analysis, and marginal rates of substitution between attributes can be estimated. As with CVM, a major drawback of this method is that it does not entail simultaneous evaluations of a number of alternatives or a choice between alternatives. On the other hand, conjoint ranking

---

[1] In the Marketing literatures, these combinations of attributes are referred to as "profiles".

[2] Several literatures include SCM as one approach of conjoint analysis. However, due to the differences between judgment data (from conjoint rating and ranking) and choice data (form stated choice method), we define SCM as an additional method other than conjoint analysis in this study. For further discussion in this issue, see Louviere (1988).

involves respondents ranking multiple sets of a number of alternative options. The response data allow the estimation of the marginal rates of substitution between the different option attributes. Since the alternatives are evaluated simultaneously, this method is closer to SCM than is conjoint rating. However, similar to the conjoint rating case, choices are not observed directly. Instead, choices between alternative options are inferred from the ranking data. Due to this essential difference between SCM and conjoint rating and ranking, Louviere and Timmermans (1990) classified rating and ranking as "stated preference models" as opposed to "stated choice models". One more difference between SCM and conjoint rating and ranking is the theoretical basis with respects to consumer behavior. SCM has a firm theoretical foundation in Random Utility Theory (RUT), which will be discussed in the next section, unlike conjoint rating and ranking.

Stated choice method has been applied in many different fields. It was originally developed in marketing research in the early 1970s, and have become widely used after a very good overview of the methods is provided by Green and Srinivasan (1978) (see, for instance, Louviere and Hensher 1983, Louviere and Woodworth 1983, Gensch 1985, Ben-Akiva *et al*. 1991, Batsell and Louviere 1991, Anderson and Wiley 1992, Ben-Akiva and Boccara 1995, Erdem and Swait 1998, Louviere *et al*. 1999). Stated choice method did not receive increasing attention in transport economics until the early 1980s. Since then, a number of important studies on demand forecast, route choice, mode choice, etc. have made use of this approach (see, for instance, Ben-Akiva and Lerman 1985, Berkovec and Rust 1985, Daly 1985, Hensher 1989, Hensher *et al*. 1989, Bradley and Gunn 1990,Ben-Akiva *et al*. 1993, Hensher 1994, Swait *et al.* 1994, Bhat 1995, Yai *et al.* 1997,Hensher 1998, Hensher 2001, Hensher and Greene 2002, Greene and Hensher 2003, Koppleman and Sethi 2005). Stepping into 1990s, economists started to apply stated choice method on the field of environmental valuation, although until now there are not many literatures as in marketing and transport studies (see, for instance, Adamowicz *et al*. 1994, Hausman *et al*. 1995, Boxall *et al*. 1996, Hanley *et al*. 1998, Blamey *et al.* 2000)

The remainder of this paper is organized as follows: Section 2 describes SCM and random utility theory, Section 3 introduces some advanced models in stated choice modeling, based on the degrees of relaxation of the Independently and Identically Distributed (IID) assumption on error terms, Section 4 discusses the procedure of a stated choice study, and Section 5 concludes the paper.

2. Stated choice method and random utility theory

Stated choice model is based on random utility theory. The basic assumption embodied in the random utility approach to choice modeling is that decision makers are utility maximizers, i.e., given a set of alternatives the decision maker will choose the alternative that maximizes his/her utility. Since the utility of an alternative for an individual U cannot be observed, it is assumed to consist of a deterministic component V and a random error term ε. Formally, individual q's utility of alternative i can be expresses as:

$$U_{iq} = V_{iq} + \varepsilon_{iq} \tag{1}$$

Hence the probability that individual q chooses alternative i from a particular set J, which comprises j alternatives, can be written as:

$$P_{iq} = P(U_{iq} > U_{jq}; \forall \quad i \neq j \in J) = P(\varepsilon_{jq} < \varepsilon_{iq} + V_{iq} - V_{jq}; \forall \quad i \neq j \in J) \tag{2}$$

To transform the random utility model into a choice model, certain assumption about the joint distribution of the vector of random error terms are required. If the random error terms are assumed to follow the extreme value type I (EV1) distribution[3] and be independently and identically distributed (IID) across alternatives and cases (or observations), the multinomial (or conditional) logit (MNL) model (McFadden 1974) is obtained. In the MNL model, the choice probability in Equation (2) is expressed as:

$$P_{iq} = \exp(\mu V_{iq}) / \sum_{j=1}^{J} \exp(\mu V_{jq}) \tag{3}$$

Then, making further assumption for the deterministic component of utility to be linear in parameters, $V_{iq} = \beta' X_{iq}$, the probability in Equation (3) is given as:

$$P_{iq} = \exp(\mu \beta' X_{iq}) / \sum_{j=1}^{J} \exp(\mu \beta' X_{jq}) \tag{4}$$

where μ represents a scale parameter that determines the scale of the utilities, which is proportional to the inverse of the distribution of the error terms. It is typically normalized to 1 in MNL model. X_{iq} are explanatory variables of V_{iq}, normally including alternative-specific constants (ASCs), the attributes of the alternative i and the social-economic characteristics of the individual q, β' is the parameter vector associated with the vector X_{iq}.

The attributes enter the utility functions at the various levels at which they are

[3] Historically, EV1 distribution has been referred to by a number of names, including Weibull, Gumbel and double-exponential.

specified in the experimental design. Including only ASCs and attributes is sufficient if individuals have homogeneous preferences. However, it is possible and frequently necessary to capture preference heterogeneity in the model by interacting respondents' socio-economic characteristics with the choice attributes or the ASCs. This involves multiplying them by either the choice attributes, which makes them attribute-specific, or by the ASCs, which makes them alternative-specific.

An important assumption of the MNL model is the independence of irrelevant alternatives (IIA) property. This property, which follows from the independence component of the IID assumption, implies that the relative choice probabilities between any two alternatives of choice set J are not affected by the inclusion or exclusion of other alternatives in that set. The IIA property is a strict assumption of the MNL model and a "reasonable approximation of more complex relationships" (Ben-Akiva and Lerman 1985). A test has been developed by Hausman and McFadden (1984) for testing the validity of the IIA assumption[4]. It is to say, if the IIA property is violated, estimating the choice model by MNL specification which exhibits IIA assumption will lead to biased estimates, therefore necessitating other model specifications. There are several advanced models that have been developed to relax the IIA assumption, which will be discussed in the next section.

3. Some models relaxing the IID assumption[5]

Substantial progress has been made in stated choice modeling, primarily through the relaxation of one or more dimensions of the IID assumption of the MNL model, resulting in more flexible model structures. In general, the additional flexibility of these advanced models comes at the cost of increased computational burden, and in some cases losing the mathematically amenable closed-form structure. These models include Nested Logit (NL) model, Heteroscedastic Extreme Value (HEV) model, Covariance Heterogeneous Nested Logit (COVNL) model, Random Parameters Logit (RPL) or Mixed Logit (ML) model, and Latent Class Logit (LCL) model[6].

---

[4] For more details on this issue, see Hausman and McFadden (1984), Louviere *et al.* (2000), Greene (2003), etc.

[5] This section is based on various sources including Bhat (1995), Allenby and Ginter (1995), Bhat (1997), Revelt and Train (1998), Louviere *et al.* (2000), Greene and Hensher (2000), McFadden and Train (2000), Hensher and Greene (2002), and Greene (2003).

[6] The Multinomial Probit (MNP) model can also be considered as a natural alternative to eliminate the IIA restriction by allowing the errors being correlated across

3.1 Nested Logit model

One way to relax the homoscedasticity assumption in the MNL model is to group the alternatives into subsets that allow the variance to differ across the subsets while maintaining the IIA assumption within the subsets. This specification defines a Nested Logit (NL) model. A primary role for NL is to allow the variances of the random components of utility to vary across subsets of alternatives (subject to the overall variance of unobserved random components of all alternatives being constant) (Louviere *et al.* 2000).

To derive the mathematical form of the model, consider a two-level NL structure[7]. Suppose an individual faces a choice of branches indexed $i = 1,2,...,I$ and elemental alternatives indexed $j = 1,2,...,J_i$ in branch i. The choice probability of alternative j in branch i by individual q can be expressed as:

$$P_{ijq} = P_{j|i,q} \cdot P_{iq} \tag{5}$$

The conditional probability $P_{j|i,q}$ can be given as:

$$P_{j|i,q} = \exp(\mu\beta' X_{ijq}) / \sum_{j=1}^{J_i} \exp(\mu\beta' X_{ijq}) \tag{6}$$

and

$$\begin{aligned} P_{iq} &= \sum_{j}^{J_i} \exp(\lambda(\alpha' Y_{iq} + \mu\beta' X_{ijq})) / \sum_{i=1}^{I}\sum_{j}^{J_i} \exp(\lambda(\alpha' Y_{iq} + \mu\beta' X_{ijq})) \\ &= \exp(\lambda\alpha' Y_{iq}) \sum_{j=1}^{J_i} \exp(\lambda\mu\beta' X_{ijq}) / \sum_{i=1}^{I} \exp(\lambda\alpha' Y_{iq}) \sum_{i=1}^{I}\sum_{j=1}^{J_i} \exp(\lambda\mu\beta' X_{ijq}) \end{aligned} \tag{7}$$

where X_{ijq} is the vector of attributes that vary with both branch and elemental levels. Y_{iq} is the vector of attributes that vary only with branch level. α' and β' are vectors of unknown parameters. λ and μ are scales parameters for branch and elemental levels, respectively. Define an Inclusive Value (IV)[8] for the i th branch as

---

alternatives and/or observations. However, due to the reason that the MNP model is equivalent in form to the RPL or ML model with certain restrictions (which will be discussed in section 3.4) on the latter model, therefore we omit a review on the MNP model and focus discussions on the other advanced models in a logit family including RPL/ML model. For more details on the MNP model, see Hausman and Wise (1978), Maddala (1983), McFadden (1989), Keane (1994), Louviere *et al.* (2000).

[7] The extension of a two-level NL structure to three-level or four-level ones can be done with the same methodology used in this paper. See Maddala (1983), Louviere (2000), Hensher and Greene (2002) for more details on the issue of three-level tree structure NL model.

[8] Inclusive Value is also termed as logsum or expected maximum utility.

$$IV_{iq} = \log(\sum_{j=1}^{J_i} \exp(\mu\beta' X_{ijq})) \quad (8)$$

Then we can rewrite Equations (6) and (7) as

$$P_{j|i,q} = \exp(\mu\beta' X_{ijq}) / \exp(IV_{iq}) \quad (9)$$

$$P_{iq} = \exp(\lambda(\alpha' Y_{iq} + IV_{iq})) / \sum_{i=1}^{I} \exp(\lambda(\alpha' Y_{iq} + IV_{iq})) \quad (10)$$

Note that the scale parameter λ associated with the branch level is often normalized to be 1. Then the scale parameters for the elemental level are left to be estimated. Alternatively, one could set $\mu = 1$ and allow λ to be estimated[9].

The IV parameter plays an important role in the NL model. It is often interpreted as a measure of dissimilarity, capturing correlations among unobserved components of alternatives in the partition. This correlation supports the claim that NL provides relaxation of independence (for alternatives sharing a partition) as well as the identical distribution assumption between alternatives in different partitions (Louviere *et al.* 2000).

3.2 Heteroscedastic Extreme Value model

The Heteroscedastic Extreme Value (HEV) model developed by Bhat (1995) and Allenby and Ginter (1995) allows different scale parameters for all alternatives in a choice set. This model is based on the same random utility structure as before and simply relaxes the assumption of equal variances. A nested logit model with a unique inclusive value parameter for each alternative (with one arbitrarily chosen variance to 1 for identification) is equivalent to an HEV specification. In mathematical term, the choice probability of alternative i from a choice set J by individual q is expressed as

$$P_{iq} = \exp(\mu_i \beta' X_{iq}) / \sum_{j=1}^{J} \exp(\mu_j \beta' X_{jq}) \quad (11)$$

where μ_i denotes the different scale parameters across alternatives. X_{iq} are explanatory variables including alternative-specific constants and the attributes of the alternative i and the social-economic characteristics of the individual q. β' is the parameter vector associated with the vector X_{iq}.

The HEV model avoids the pitfalls of the IID property by allowing different scale

[9] Hensher and Greene (2002) mention that the impact of normalization on the scales of some parameters may produce internal inconsistency of the model if not handled properly. Typically, if the same parameter appears in several nests, normalization from the bottom may cause problems, as the parameter will be scaled differently in each nest.

parameters across alternatives. Intuitively, we can explain this by realizing that the random term represents unobserved attributes of an alternative; that is, it represents uncertainty associated with the expected utility (or the observed part of utility) of an alternative. The scale parameter of the error term, therefore, represents the level of uncertainty (the lower the scale, the higher the uncertainty) (Louviere *et al.* 2000).

3.3 Covariance Heterogeneous Nested Logit model

Bhat (1997) proposed a modification to the nested logit model that allows heterogeneity across individuals in the covariance of nested alternatives, termed as Covariance Heterogeneous Nested Logit (COVNL) model. As an alternative specification of NL logit and HEV models, COVNL model estimates a model in which the similar scale parameters across alternatives are a function of individual-specific and/or alternative-specific variables as sources of scale decomposition. Mathematically, the function of similar scale parameter μ_i is given as:

$$\mu_i = F(\psi + \gamma' Z_{iq}) \tag{12}$$

where Z_{iq} is a vector of individual and/or alternative related characteristics, ψ and γ' are parameters to be estimated, and F is a transformation function that ensures μ_i is bounded between 0 and 1[10]. Then, COVNL choice probabilities are given by Equation (13), while μ_{iq} is given by Equation (12):

$$P_{iq} = \exp(\mu_i \beta' X_{iq}) / \sum_{j=1}^{J} \exp(\mu_i \beta' X_{jq}) \tag{13}$$

If $\gamma' = 0$ in Equation (12), covariance heterogeneity is absent and the COVNL reduces to a NL model. The COVNL model is more complex than the simple NL model but still retains a closed form structure. Due to both the introduction of additional variables and the incorporation of the covariance structure, this model is statistically and behaviorally superior to the corresponding NL and HEV models. For example, the NL and HEV models can partially capture the heteroscedasticity by specifications of the scale parameters, however, the origin of the variability would not be explicit without formulating a covariance structure of scale parameters as COVNL does.

3.4 Random Parameters Logit or Mixed Logit model

Other than HEV and COVNL models, the Random Parameters Logit (RPL) model

---

[10] McFadden (1981) noted that a global sufficiency condition of the nested choice model to be consistent with random utility maximization is that the parameters of inclusive value be in the 0-1 range.

(also be termed as Mixed Logit (ML) model) allows for a more heightened level of flexibility by specifying taste coefficients to be randomly distributed across individuals (see Revelt and Train 1998, McFadden and Train 2000, Louviere *et al.* 2000). Additionally, RPL/ML model has a considerable advantage not available in any of the other models mentioned above. It is that RPL/ML model can account for potential correlation over repeated choices made by each individual, although imposing a first-order autoregressive (AR1) process is extremely complex.

The model is a generalization of the MNL model, summarized as below:

$$P_{iqt} = \exp(\alpha' + \beta' X_{iqt} + \varphi' F_{iqt}) / \sum_{j=1}^{J} \exp(\alpha' + \beta' X_{jqt} + \varphi' F_{jqt}) \tag{14}$$

where

- α' is a vector of fixed or random alternative-specific constant associated with $i = 1,..., J$ alternatives and $q = 1,..., Q$ individuals, and one of these ASCs should be identified as 0.
- β' is a parameter vector that is randomly distributed across individuals.
- φ' is a vector of non-random parameters.
- X_{iqt} is a vector of individual-specific characteristics and alternative-specific attributes at observation t, and is estimated with random parameters.
- F_{iqt} is a vector of individual-specific characteristics and alternative-specific attributes at observation t, and is estimated with fixed parameters.

In this specification, a subset or all of α' and the parameters in the β' vector can be assumed to be randomly distributed across individuals[11]. These random parameters can then be defined as a function of characteristics of individuals and/or other attributes that are choice invariant. Based on these defined attributes, the mean and standard deviations of specified random parameters and contributions from these choice invariant attributes on random parameters are estimated by using Maximum Simulated Likelihood (MSL) method. The RPL/ML model is sufficiently flexible that it provides the modeler a tremendous range within which to specify individual unobserved heterogeneity. To some extent, this flexibility offsets the specificity of the distributional assumptions (Greene and Hensher 2003).

A further important issue should be mentioned here is that the RPL or ML model is equivalent in form to the Multinomial Probit (MNP) model, even though the variances of the random component take on a different distribution (i.e., EV1 compared to normal),

[11] The distributions of random parameters can be considered, for example, normal distribution, lognormal distribution, and triangular distribution, etc.

if we assume (a) that the alternative-specific constants are random, (b) choice invariant characteristics variables that produce individual heterogeneity in the averages of the randomly distributed parameters are excluded, and (c) that the full (i.e., including the variances) lower triangular matrix of covariance is unrestricted. This equivalence is very important, since this special case of the RPL or ML model provides an alternative method of estimation to MNP (*Louviere et al.* 2000, ch. 6).

3.5 Latent Class Logit Model

The Latent Class Logit (LCL) model, unlike RPL/ML model which specifies the random parameters to follow a continuous joint distribution, assumes that a discrete number of classes are sufficient to describe the joint function of the parameters. Therefore, the unobserved heterogeneity is captured by these latent classes in the population, each of which is associated with a different parameter vector in the corresponding utility. The LCL has often been used in marketing research instead of RPL/ML model, while there are few studies in other fields such like transportation and environmental valuation.

The choice probability of individual q of class s is expressed as:

$$P_{iq|s} = \exp(\mu_s \beta_s' X_{iq}) / \sum_{j=1}^{J} \exp(\mu_s \beta_s' X_{jq}) \qquad s = 1,\ldots,S \tag{15}$$

which is a simple MNL specification in class s. Additionally, one can construct a classification model as a function of some individual-specific attributes to explain the heterogeneity across classes. The LCL model simultaneously estimates Equation (15) for S classes and predicts the probability H_{qs} as individual q being in class s. Then, the unconditional probability of choosing alternative i is given as:

$$P_{iq} = \sum_{s=1}^{S} P_{iq|s} H_{qs} \tag{16}$$

An issue to be noted is the choice of S, the number of classes. Since this is not a parameter, hypotheses on S cannot be tested directly. However, as Louviere *et al.* (2000, ch. 10) mentioned that a number of methods to decide S have been used based on the Akaike Information Criterion (AIC) and its variants. AIC and Consistent AIC (CAIC), which are given in Equations (17) and (18), are used to guide model selection.

$$AIC = -2(LL(\hat{\theta}) - S \cdot K_s - (S-1)K_c) \tag{17}$$

$$CAIC = -2(LL(\hat{\theta}) - (S \cdot K_s + (S-1)K_c - 1)(\ln(2N)+1) \tag{18}$$

where $LL(\hat{\theta})$ is the log likelihood at the estimated parameters $\hat{\theta}$, K_s is the number of elements in the utility function of the class-specific choice models, K_c is the total number of parameters in the classification model, and N is the number of observations in the sample. The value of S that minimizes each of the measures of AIC and CAIC suggests which model should be preferred (Louviere *et al.* 2000, ch. 10).

4. Procedure of a Stated Choice study

A Stated Choice (SC) study comprises a number of stages. Many literatures introduce the common steps for the SC practices (see, for instance, Louviere and Timmermans 1990, Hensher 1994, Louviere *et al.* 2000). The steps suggested in these references can be grouped into four broad stages: questionnaire design, data collection, model estimation and assessment, and application of model results.

4.1 Questionnaire design

The design of the survey instrument is crucial for the quality of the survey results. A number of important issues should be identified. These are the selection of choice attributes and levels, the experimental design, and presentation of choice tasks.

Selection of choice attributes and levels

The first task in the choice context is to select the set of choice attributes which are likely to be most important. The set normally includes those which are commonly found to be important (for example, cost, access time, in-vehicle time, etc. in a transportation mode choice analysis), plus any instrumental factors in the policies or scenarios to be studied. Sometimes, however, the set of choice attributes important to the respondent is not the same as might be deduced from most existing models.

It is noted that different choice attributes may be important to different people, so that, for some, important attributes are missing from the experiment. One may consider adding attributes as many as accountable into the choice experiment. However, as the number of attributes increase, task complexity increases because of the number of things to which respondents must attend. Several studies have shown that task complexity leads to preference instability as the experiment progresses (see, for instance, Hensher 2004, etc.). As task complexity of a choice experiment increases, respondents will find more and more difficult to finish the questionnaires. The cognitive burden on respondents is very much determined by the topic of the study at hand. Carson *et al.* (1994) state that the average number of attributes included in questionnaires is around seven. Generally, an investigation into individual choices related to familiar decisions allows a greater number of attributes to be evaluated compared with a more unfamiliar choice situation. On the other hand, if something is missing which is very important to the respondent and which affects the credibility of the other variables, the results may also be less valid. Therefore, from this view, the pilot test for examining the validness of choice attributes is necessary.

Issues involved in the selection of the levels of the chosen attributes include the range and the measurement of the attributes. With respect to the range, an important consideration is the current range experienced by respondents. A often used design is to identify these ranges at both the extreme high and low ends, as Louviere *et al.* (2000) suggest that the wider the range of levels, the more likely it will be that more subjects agree that some levels are "high" whereas others are "low". However, it is most important to detect that these extreme levels are realistic and acceptable by the respondents. An excessively limited range for the attributes hamstrings subsequent analyses if respondents find them unbelievable.

With respect to the measurement of the attributes, a distinction is made between subjective and objective attributes. The latter can be objectively defined, for instance in terms of distance or dollar amounts. Subjective attributes, such as environmental quality, are more difficult to be defined. An ordinal scale of high, medium and low for an attribute like cleanliness at a destination is suggested in literatures. However, for this kind of ordinal definition, a careful description of each level is required to make them be understandable by respondents.

Experimental design

After the choice attributes and their level are determined, and the choice alternatives are selected, developing an experimental design is sequentially the next step. The experimental design produces the choice sets for the Stated Choice Model questionnaire to enable the estimation of the contributions of the attributes to the utility function of the alternatives.

A full factorial design is most ideal due to the reason that it enables the estimation of the main effects and all attribute interactions. However, as the number of attributes and/or their levels increase, the number of choice sets increase dramatically. For instance, if there are six attributes each defined at three levels, the total number of choice sets is 729 (3^6). Clearly, it would be impossible for respondents to assess all choice sets. Hence, a fractional factorial design is often used. Fractional factorials are generally orthogonal and allow the estimation of at least all main effects. Some or all two-way or high orders interactions can also be estimated by applying fractional factorial designs. However, since this is done based on increasing the number of choice sets in the design, the task complexity would also increase. Therefore, the selection of the experimental design by the analyst involves a consideration of the trade-off between cognitive complexity and analytical sophistication. Carson *et al.* (1994) note that in most studies respondents evaluate between one and sixteen choice sets, with the average being somewhere around eight choice scenarios per respondent.

In some cases, the number of choice sets may still be too demanding for survey respondents even after fractional factorial design. To deal with this, the design is commonly divided into subsets. This procedure is refereed as “blocking” in the design literature. It therefore creates a number of different versions of questionnaires with each respondent exposed to one version only. Blocking can be done randomly or in a systematic fashion.

Presentation of questionnaires

The presentation of the SCM questionnaires can range from hard copy to various multi-media modes. The choice of which mode to use is often determined by budget constraints. Regardless of the mode selected, the overall presentation of the questionnaire requires careful consideration. The objective is to present the choice experiment as an approximation of actual choice situations. To that end, background information needs to be provided. It is important that this information is consistent with the information that respondents normally have in order to make actual choices. In other words, the frame must be appropriate to the decision context.

The approximation of actual situations in the questionnaire, and the manner in which the questionnaire is presented requires careful testing before the questionnaire is put in front of the survey respondents. Conducting focus group is an usual way for exploratory researches. Focus groups are conducted to gain an understanding about the research issue by obtaining feedback from the target audience in a small group environment. Focus groups can also be used to test draft questionnaires on cognitive issues.

4.2 Data collection

Data collection mainly involves the issues on sampling strategy and response collection method.

Sampling strategy

The choice of survey population obviously depends on the objective of the survey. Given the survey population, a sampling strategy has to be determined. Possible strategies include a simple random sample, a stratified random sample or a choice-based sample. A simple random sample is generally a reasonable choice. One reason for choosing a more specific sampling method may be the existence of a relatively small but important sub-group which is of particular interest to the study. Another reason may be to increase the precision of the estimates for a particular sub-group. In practice the selection of sample strategy and sample size is also largely dependent on the budget available for the survey (Alpizar *et al.* 2003).

Louviere *et al.* (2000) provide a formula to calculate the minimum sample size. The size of the sample, n, is determined by the desired level of accuracy of the estimated probabilities, $\hat{p}$. Let p be a true proportion of the relevant population, a is the percentage of deviation between $\hat{p}$ and p that can be accepted and α is the confidence level of the estimation such that: $\Pr(|\hat{p}-p| \leq ap) \geq \alpha$ for a given n. Given this, the minimum sample size is defined as:

$$n \geq \frac{1-p}{pa^2}\Phi^{-1}\left(\frac{1+\alpha}{2}\right) \tag{19}$$

where Φ is the standard normal Cumulative Distribution Function (CDF). Note that n refers to the size of the sample and not the number of observations. Since each individual makes r succession of choices in a choice experiment, the number of observations will be much larger (a sample of 400 individuals answering 8 choice sets each will result in 3200 observations). From this view, one of the advantage of choice experiments is that the amount of information extracted from a given sample size is much larger than, for example, using referendum-based methods and, hence, the efficiency of the estimates is improved. The formula in Equation (19) is only valid for a simple random sample and with independence between the choices[12].

Response collection method

Once the questionnaire is designed and tested, the sampling strategy is decided and the sample size is calculated, the survey can be put into the field. The selection of response collection methods depends on the type of respondents, the complexity of the choice decision or of the product being studied, and the budget available for the study.

A singularly cost-effective method of response collection is the mail survey, which is most effective when respondents can be recruited by telephone or other means[13].

[12] For more details on this issue, see Ben-Akiva and Lerman (1985).

[13] A similar method with mail survey, termed as posting, can be also used in some simple and familiar products' choice experiments. It is executed by delivering questionnaires into the respondents' posts and asking the respondents to mail them

However, mail survey suffers the problems such as low response rate, and relative more invalid answers compared with other methods due to the cognitive difficulty on survey. A telephone survey often used in other surveys is seldom considered as an effective method in stated choice experiments, since interviewers are quite difficult to explain the choice sets to the respondents by telephone. However, a mail survey combined with an advance telephone call and/or a follow-up telephone call is often used to raise up the response rates. In addition, in some cases, face-to-face interview may be applied. Although personal interview has a number of advantages compared with mail survey, it may be very expensive to execute depending on sample sizes.

Recently, computer-based interviews are developed very fast. They come in several forms: (1) a self-completion survey is sent to respondents on floppy disk or CD-ROM, and mailed back to the research upon completion; (2) personal interviews are conducted using a computer, with interviewers and/or respondents keying in responses to questions. The former method is often more useful in business-to-business applications, whereas the latter is more often used for interviewing consumers. Computer-based interview has the advantages of flexibility (i.e., questionnaire flow can be altered in real time) and improved data quality (i.e., error checking occurs at the time of response) (Louviere *et al.* 2000, ch. 9).

4.3 Model estimation and assessment

To analyze the response data, a statistical choice model is required. As discussed in Section 2 and Section 3, different models such as MNL, NL, HEV, COVNL, RPL/ML, and LCL, are obtained from different error term assumptions and can be used for estimation. To assess that one model is more superior to another one, two tests for nested and non-nested models are introduced as follows[14].

The most common test undertaken to compare any two nested models is the likelihood ratio (LR) test. The formula of LR statistics is given as:

$$LR = -2(LL1 - LL2) \tag{20}$$

where $LL1$ and $LL2$ are the log likelihood at convergence for model 1 and model 2 using same data set. Define n as the difference in the degrees of freedom for two models. The calculated LR is compared to the critical value from a chi-squared test table at an appropriate level of statistical significance (e.g. 0.05 being the most used level in literatures) for the number of degrees of freedom n. If LR is greater than the critical

back upon completion.

[14] A test for choice-based samples for nested and non-nested models is not introduced in this review. A detailed discussion on this issue can be found in the study of Louviere *et al.* (2000).

value, then we can conclude that the two models are statistically different, rejecting the null hypothesis of no difference.

A test for non-nested choice models based on the Akaike Information Criterion (AIC) has been proposed by Ben-Akiva and Swait (1986). Suppose model 1 explains choices using K_1 variables, while model 2 explains the same choices using K_2 variables; assume that $K_1 \geq K_2$ and either the two models have different functional forms or the two sets of variables are different by at least one element. Define the fitness measure for model j, j=1, 2:

$$\rho_j^2 = 1 - \frac{L_j - K_j}{L(0)} \tag{21}$$

where L_j is the log likelihood at convergence for model j and $L(0)$ is the log likelihood for constants only. Ben-Akiva and Swait (1986) show that under the null hypothesis that model 2 is the true model, the probability that the fitness measure in Equation (21) for model 1 will be greater than that of model 2 is asympototically bounded by a function given in Equation (22):

$$\Pr(|\rho_2^2 - \rho_1^2| \geq Z) \leq \Phi(-\sqrt{-2ZL(0) + (K_1 - K_2)}) \tag{22}$$

where Z is the difference of the fitness measures between model 1 and model 2 and assumed larger than zero, Φ is the standard normal Cumulative Distribution Function (CDF). Therefore, Equation (22) sets an upper bound for the probability that one incorrectly selects model 1 as the true model if model 2 is the true model.

4.4 Application of model results

Once the statistical choice model has been estimated and assessed, the model output is applied to the research problem at hand. The estimated coefficients of the explanatory variables are interpreted in the context of the study. The marginal contribution to the utility of the alternatives can be compared for the different attributes included in the model. The influence of individual socio-economic characteristics on choice utilities can also be assessed.

One approach of application of the model results is using the relevant non-price attribute level coefficients and price coefficients to estimate the implicit price change in the non-price attributes. For instance, regarding the transport economics, one important policy analysis item, the value of time saving, can be calculated by the ratio of the estimated time and cost coefficients. Examples in this application can be found in recent studies such as Louivere *et al.* (2000, ch. 11), Lam and Small (2001), Hensher (2001), Greene and Hensher (2003), Hensher (2004), Brownstone and Small (2005), and

Hess *et al.* (2005) etc.

A further application of the model results elasticity analysis, which expresses the percentage change in a response (e.g., market share) caused by a 1 per cent change in a certain variable (e.g., price). Such applications can be found, for instance, in Louivere *et al.* (2000, ch. 11), Ortuzar and Gonzalez (2002), Greene and Hensher (2003), and Menendez *et al.* (2004) etc.

Finally, for some fields like transportation planning and environmental economics, which both deal with public policy, the Compensating Variation (CV) can be calculated by application of the model results for welfare analysis. See, for instance, Louivere *et al.* (2000, ch. 12), Alpizar *et al.* (2003), and Li *et al.* (2004) etc.

5. Conclusion

This paper has reviewed Stated Choice Method. Attentions have been focused on the theoretical background of SCM, the application and empirical modeling of SCM, the design of stated choice experiments, and the procedure of SCM. Stated Choice Method, as one of the stated preference modeling techniques, has been developed to be capable of analyzing a number of possible choice situations by requiring variability of choice attributes through the use of an appropriate experimental design. Further, in comparison with the other stated preference models, a SCM application simulates more closely actual choice behavior and is firmly grounded in the behavioral foundations of random utility theory. As a result, SCM has become an important method for analyzing various policy impacts.

However, a good method does not necessarily ensure that an application of it would be always successful. To succeed in an application of stated choice method, analysts should carefully deal with every step of the procedure. In most cases, failure of one step may lead to unsuccessfulness of the project, since the failure usually cannot be compensated by other steps. For example, as for the issue on sampling and design, a good design will not compensate inadequate sampling, and vice versa. Therefore, as a topic for future research, a kind of guideline for SCM such like NOAA guideline for CVM (Contingent Valuation Method) is worthy of being developed.

References

Adamowicz W. and Swait J. D. (1994), "Combining stated and revealed preference methods for valuing environmental amenities". *Journal of Environmental Economics and Management* 26: 271-292.

Allenby G. and Ginter J. (1995), "The effects of in-store displays and feature advertising on consideration sets". *International Journal of Research in Marketing* 12: 67-80.

Alpizar F., Carlsson F. and Martinsson P. (2003,) "Using choice experiments for non-market valuation". *Economic Issues* 8(1): 83-110.

Anderson D. A. and Wiley J. B. (1992), "Efficient choice set designs for estimating availability cross-effects models". *Marketing Letters* 3: 357-370.

Batsell R. R. and Louviere J. J. (1991), "Experimental analysis of choice". *Marketing Letters* 2: 199-214.

Ben-Akiva M. E. and Boccara B. (1995), "Discrete choice models with latent choice sets". *International Journal of Research in Marketing* 12(1): 9-24.

Ben-Akiva M. E., Bolduc D. and Bradley M. (1993), "Estimation of travel choice models with randomly distributed values of time". *Transportation Research Record* 1413: 88-97.

Ben-Akiva M. E. and Lerman S. (1985), *Discrete choice analysis: theory and application to travel demand*, Cambridge, Mass: MIT Press.

Ben-Akiva M. E., Morikawa T. and Shiroishi F. (1991), "Analysis of the reliability of preference ranking data". *Journal of Business Research* 23: 253-268.

Ben-Akiva M. E. and Swait J. D. (1986), "The akaike likelihood ratio index". *Transportation Science* 20(2): 133-136.

Berkovec J. and Rust J. (1985), "A nested logit model of automobile holding for one vehicle households". *Transportation Research* 19B(4): 275-286.

Bhat C. R. (1995), " A heteroscedastic extreme value model of intercity travel mode choice". *Transportation Research* 29B(6): 471-483.

Bhat. C. R. (1997), "A nested logit model with covariance heterogeneity". *Transportation Research* Part B 31: 11-21.

Bhat C. R. (2003), "Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences". *Transportation Research* Part B 37: 837-855.

Bhat C. R. and Gossen R. (2004), "A mixed multinomial logit model analysis of weekend recreational episode type choice". *Transportation Research* Part B 38: 767-787.

Blamey R. K., Bennett J. W., Louviere J. J. Morrison M. D. and Rolfe J. (2000), "A test of policy labels in environmental choice modeling studies". *Ecological Economics* 32:

269-286.

Boxall P., Adamowicz W., Williams M., Swait J. D. and Louviere J. J. (1996), "A comparison of stated preference approaches to the measurement of environmental values". *Ecological Economics* 18: 243-253.

Bradley M. A. and Gunn H. (1990), "Stated preference analysis of values of travel time in the Netherlands". *Transportation Research Record* 1285: 78-89.

Brownstone D. and Small K. A. (2005), "Valuing time and reliability: assessing the evidence from road pricing demonstrations". *Transportation Research* Part A 39: 279-293.

Carson R., Louviere J. J., Anderson D., Arabie P., Bunch D., Hensher D., Johnson R., Kuhfeld W., Steinberg D., Swait J. D., Timmermans H. and Wiley J. (1994), "Experimental analysis of choice". *Marketing Letters* 5(4): 351-368.

Daly A. J. (1985), "Estimating "tree" logit models". *Transportation Research* 21B(4): 251-268.

Erdem T. and Swait J. D. (1998), "Brandy equity as a signaling phenomenon". *Journal of Consumer Psychology* 7(2): 131-157.

Gensch D. (1985), "Empirically testing a disaggregate choice model for segments". *Journal of Marketing research* 22: 462-467.

Green P. E. and Srinivasan V. (1978), "Conjoint analysis in consumer research: issues and outlook". *Journal of Consumer Research* 5:103-123.

Greene W. H. (2003), *Econometric analysis*, Fifth edition by Pearson Education International, New York University.

Greene W. H. and Hensher D. A. (2003), "A latent class model for discrete choice analysis: contrasts with mixed logit". *Transportation Research* Part B 37: 681-698.

Hanley N., Wright R. E. and Adamowicz V. (1998), "Using choice experiments to value the environment: design issues, current experience and future prospects". *Environmental and Resource Economics* 11(3-4): 413-428.

Hausman J. A., Leonard G. K. and McFadden D. (1995), "A utility-consistent combined discrete choice and count data model assessing recreational use losses due to natural resource damage". *Journal of Public Economics* 56: 1-30.

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

Hausman J. A. and Wise D. A. (1978), "A conditional probit model for qualitative choice: discrete decisions recognizing interdependence and heterogeneous preferences". *Econometrica* 46: 403-426.

Hensher D. A. (1989), "Behavioral and resource values of travel time savings: a bicentennial update". *Australian Road Research* 19(3): 223-229.

Hensher D. A. (1994), "Stated preference analysis of travel choices: the state of the practice". *Transportation* 21: 107-133.

Hensher D. A. (1998), "Establishing a fare elasticity regime for urban passenger transport: non-concession commuters". *Journal of Transport Economics and Policy* 32(2): 221-246.

Hensher D. A. (2001), "The sensitivity of the valuation of travel time saving to the specification of unobserved effects". *Transportation Research* part E 37: 129-142.

Hensher D. A. (2004), "Identifying the influence of stated choice design dimensionality on willingness to pay for travel time saving". *Journal of Transport Economics and Policy* 38(3): 425-446.

Hensher D. A., Barnard P., Milthorpe F. and Smith N. (1989), "Urban tollways and the valuation of travel time savings". *Economic Record* 66(193): 146-156.

Hensher D. A. and Greene W. H. (2002), "Specification and estimation of the nested logit model: alternative normalizations". *Transportation Research* Part B 36: 1-17.

Hess S., Bierlaire M. and Polak J. W. (2005), "Estimation of value of travel-time saving using mixed logit models". *Transportation Research* Part A 39: 221-236

Keane M. (1994), "Modeling heterogeneity and state dependence in consumer choice behavior". Working paper, Department of Economics, University of Minnesota.

Koppleman F. S. and Sethi V. (2005), "Incorporating variance and covariance heterogeneity in the Generalized Nested Logit model: an application to modeling long distance travel choice behavior". *Transportation Research* Part B 39: 825-853.

Kores E. P. and Sheldon R. J. (1988), "Stated preference methods: an introduction". *Journal of Transport Economics and Policy.* Volume XXII No. 1: 11-25.

Lam T. C. and Small K. A. (2001), "The value of time and reliability: measurement from a value piecing experiment". *Transport Research part* E 37: 231-251

Li C. Z., Kuuluvainen J., Pouta E., Rekola M. and Tahvonen O. (2004), "Using choice experiments to value the Natura 2000 nature conservation programs in Finland". *Environmental & Resource Economics* 29: 361-374

Louviere J. D. (1988), "Conjoint analysis modeling of stated preference: a review of theory, methods, recent developments and external validity". *Journal of Transport Economics and Policy.* Volume XXII No. 1: 93-119.

Louviere J. J. and Hensher D. A. (1983), " Using discrete choice models with experimental design data to forecast consumer demand for a unique cultural event". *Journal of Consumer Research* 10(3): 348-361.

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

Louviere J. J., Meyer R. Bunch D. Carson R. Dellaert B. Hanemann W. M. Hensher D. A. and Irwin J. (1999), "Combining sources of preference data for modeling complex decision processes". *Marketing Letters* 10(3): 187-217.

Louviere J. J. and Timmermans H. (1990), "Stated preference and choice models applied to recreational research: a review". *Leisure Sciences* 12: 9-32.

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

Maddala G. S. (1983), *Limited-dependent and qualitative variables in econometrics*, Cambridge University Press.

McFadden D. (1974), *Conditional logit analysis of quanlitative choice behavior*, Zarembka P. (ed.), Frontiers in econometrics, New York: Academic Press. 105-142.

McFadden D. (1981), *Econometric models of probabilistic choice*, Manski C. and McFadden D. (eds.), Structural analysis of discrete data with econometric applications, Cambridge, Mass: MIT Press. 198-272.

McFadden D. (1989), "A method of simulated moments for estimation of discrete response models without numerical integration". *Econometrica* 57(5): 995-1026.

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

Menendez L. G., Zarzoso I. M. and Miguel D. P. (2004), "Determinants of mode choice between road and shipping for freight transport". *Journal of Transport Economics and Policy* 38(3): 447-466

Ortuzar J. D. and Gonzalez R. M. (2002), "Inter-island travel demand response with discrete choice models". *Journal of Transport Economics and Policy* 36(1): 115-138

Revelt D. and Train K. (1998), "Incentives for appliance efficiency in a competitive energy environment: random parameters logit models of households' choices". *Review of Economics and Statistics* 80(4): 647-657.

Swait J. D. and Louviere J. J. (1993), "The role of the scale parameter in the estimation and comparison of multinomial logit models". *Journal of Marketing Research* 30: 305-314.

Yai T., Iwakura S. and Morichi S. (1997), "Multinomial probit with structured covariance for route choice behavior". *Transportation Research* 31B(3): 195-207.