Representation of heteroskedasticity in discrete choice models

Marcela A. Munizaga ^{a,*}, Benjamin G. Heydecker ^b, Juan de Dios Ortúzar ^c

^a *Department of Civil Engineering, Universidad de Chile, Casilla 228-3, Santiago, Chile*

^b *Centre for Transport Studies, University College London, Gower Street, London WC1E 6BT, UK*

^c *Department of Transport Engineering, Pontificia Universidad Católica de Chile, Casilla 306, Santiago 22, Chile*

Received 15 May 1998; received in revised form 17 May 1999; accepted 18 May 1999

Abstract

The Multinomial Logit, discrete choice model of transport demand, has several restrictions when compared with the more general Multinomial Probit model. The most famous of these are that unobservable components of utilities should be mutually independent and homoskedastic. Correlation can be accommodated to a certain extent by the Hierarchical Logit model, but the problem of heteroskedasticity has received less attention in the literature. We investigate the consequences of disregarding heteroskedasticity, and make some comparisons between models that can and those that cannot represent it. These comparisons, which use synthetic data with known characteristics, are made in terms of parameter recovery and estimates of response to policy changes. The Multinomial Logit, Hierarchical Logit, Single Element Nested Logit, Heteroskedastic Extreme Value Logit and Multinomial Probit models are tested using data that are consistent with various error structures; only the last three can represent heteroskedasticity explicitly. Two different kinds of heteroskedasticity are analysed: between options and between observations. The results show that in the first case, neither the Multinomial Logit nor the Single Element Nested Logit models can be used to estimate the response to policy changes accurately, but the Hierarchical Logit model performs surprisingly well. By contrast, in a certain case of discrete heteroskedasticity between observations, the simulation results show that in terms of response to policy variations the Multinomial Logit model performs as well as the theoretically correct Single Element Nested Logit and Multinomial Probit models. Furthermore, the Multinomial Logit Model recovered all parameters of the utility function accurately in this case. We conclude that the simpler members of the Logit family appear to be fairly robust with respect to some homoskedasticity violations, but that use of the more resource-intensive Multinomial Probit model is justified for handling the case of heteroskedasticity between options. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Transport demand; Discrete choice models; Heteroskedasticity

* Corresponding author. Tel.: +56 2 6784649; fax: +56 2 6718788.

E-mail address: mamuniza@cec.uchile.cl.

1. Introduction

Although the Multinomial Logit model (MNL) is widely used in transport demand modelling, there has always been concern about its inherent property of Independence from Irrelevant Alternatives (IIA). This property has its basis in the convenient but simplistic assumption of Independent and Identically Distributed (IID) unobservable components of utility. In recent years, the estimation of Probit models has become practical though it remains substantially more computationally demanding. Thus, it is now possible to investigate whether the bias caused by the assumption of IID error terms in the MNL is an acceptable consequence of avoiding the extra computing effort entailed in estimating a Probit one.

In this paper we study the effect of heteroskedasticity (error terms with different variance), of which two different kinds are considered. The first is *between options*, a case that can arise from levels of information that differ between options (e.g. perception of regularly versus rarely selected options), differences in the variability inherent in the options, or for several other reasons. The second kind is *between observations* and this is present for example (in a binary or discrete form) in mixed data experiments where Revealed Preference (RP) and Stated Preference (SP) survey observations are both used to model the same options.

Our study of the performance of various different models is based on simulation. We identify some cases where simulation can be used to support intuition in the identification of the capabilities of each model and illustrate the magnitude of the biases that arise when inappropriate models are used. We shed some light on these points by using simulation experiments and are also able to verify empirically certain known results, such as that an MNL model is equivalent to a Probit one that has homoskedastic and un-correlated errors (IID Probit). We investigate the behaviour of various models when their respective assumptions hold and test their performance when some do not hold. We target our analysis on practical considerations, and in that sense design the experiments and consider the models according to what is needed and what is used by practitioners. Also, we prefer the more parsimonious models when several perform satisfactorily.

The remainder of the paper is organised as follows. In Section 2, we describe the covariance structure underlying the models considered and their estimation procedures; we also review briefly other models that could play a role in addressing the issues examined. Section 3 describes the simulation experiments and present its main results, making clear the scope of our work. We also make some recommendations for the interpretation of our simulation results: they bear directly on the cases of heteroskedasticity between options and discrete heteroskedasticity between observations that we simulate, and indicate areas for further investigation. Finally, our conclusions are given in Section 4. The Multinomial Probit (MNP) was the most reliable model in all circumstances. However, we find the workhorses of the Logit family (MNL and HL) surprisingly good. The MNL is remarkably robust in the case of heteroskedasticity between observations, whilst the Hierarchical or Tree Logit (HL) performs surprisingly well in the presence of heteroskedasticity between options except for very strong policy changes.

2. Forms of the Logit and Probit models

Discrete choice theory is well established and has several associated econometric tools. We summarize here some of the basic concepts involved; the interested reader can refer to Ben-Akiva

and Lerman (1985), Hensher and Johnson (1981) or Ortúzar and Willumsen (1994) for further details. The microeconomic view of the discrete-choice process is that each individual will choose the option that has for her the greatest utility. In general the utility U_{in} of option i to individual n , corresponding to the relative attractiveness of the option, may be treated as a random variable consisting of the sum of an observable (or systematic) part V_{in} plus an error term ε_{in} with zero mean and a certain distribution:

$$U_{in} = V_{in} + \varepsilon_{in}. \quad (1)$$

The error term represents the modeller's inability to observe all the variables that influence the choice decision, i.e. measurement errors, differences between individuals, the individuals' erroneous perceptions of the attributes, and the randomness inherent in human nature. The choice probability for option i can be written as:

$$P(i|C_n) = \Pr[V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn} \quad \forall j \in C_n], \quad (2)$$

where C_n is the set of options available to individual n . To derive a specific random utility model we require an assumption about the joint probability distribution of the full set of error terms. According to Ben-Akiva and Lerman (1985, p. 69) "... One logical assumption is to view the disturbances as the sum of a large number of unobserved but independent components. By the central limit theorem the distribution of the disturbances would tend to be Normal". In this case, the resulting function is called the *Probit* model, and the choice probability of option i is given by:

$$P(i|C_n) = \Pr[\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn} \quad \forall j \in C_n] = \int_{-\infty}^{V_{in}-V_{1n}} \int_{-\infty}^{V_{in}-V_{2n}} \dots \int_{-\infty}^{V_{in}-V_J} N(0, \Sigma_{\varepsilon-\varepsilon_i}) d\varepsilon, \quad (3)$$

where $N(0, \Sigma_{\varepsilon-\varepsilon_i})$ is a multivariate Normal density function with zero means and covariance matrix $\Sigma_{(\varepsilon - \varepsilon_i)}$, and J is the number of options in the set C_n . Unfortunately, this integral cannot be solved analytically so either simulation or an approximation must be used to evaluate it. This theme is discussed in Appendix A.

Because of the inconvenience of the MNP model, being both difficult to estimate and to interpret, a simpler model is often used in practice. The MNL model can be derived if the error terms are mutually independent and identically distributed Extreme Value (also called Gumbel or Weibull) random variables. It is called a "Probit-like" model by Ben-Akiva and Lerman (1985) because the Extreme Value distribution is similar to the Normal distribution. It can also be justified from the derivation of the Extreme Value distribution (Gumbel, 1957) as the distribution of the probability that the maximum of several IID variables is greater than a certain value. The MNL has the following simple form (McFadden, 1974):

$$P(i|C_n) = \frac{\exp(\mu V_{in})}{\sum_{j \in C_n} \exp(\mu V_{jn})}, \quad (4)$$

where μ is a parameter related to the variance σ^2 of the error term $\mu = \pi/\sqrt{6}\sigma$. This parameter is unidentifiable and can be set to any value without affecting the model because scaling the utility

values V_{in} can subsume its role; thus, it is usually set equal to unity. Fixing the value of the parameter μ is equivalent to scaling the covariance matrix. This is also required in the Probit model where the covariance matrix is usually scaled so that the element σ_{11}^2 in the first cell is equal to unity (see for example Bolduc, 1992).

The MNL form of the Logit model is widely used in various fields, though the IID assumption means that the error terms are supposed to be independent and homoskedastic. The assumption of homoskedasticity fails in at least two common practical cases. When using two data sets concurrently, for example one from an SP experiment and another from an RP one, it is often not reasonable to assume that the error variances of these two kinds of observations are identical. We call this heteroskedasticity *between observations*¹. Another case is exemplified by ranking SP experiments where respondents have to consider several options and rank them in order of preference. Pearmain et al. (1991), amongst others, have argued that people may be more precise when ranking options that have either high or low rank, and less precise with those that are intermediate. This is a clear example of heteroskedasticity *between options*, where the intermediate options are expected to have larger error variances. Some other formulations of Logit models can address these problems in part.

The *Hierarchical or Tree Logit* (HL) model (see Williams, 1977; Daly, 1987) can be used to model discrete choices when there is positive correlation between the error terms within each of some mutually exclusive groups of observations, but it maintains the property of homoskedasticity. The genesis of this model is the recognition that some options are perceived as being inherently similar. The positively correlated options are grouped in nests within each of which the IID assumption applies. This gives rise to a structured covariance matrix where if two options are in the same nest, the corresponding off-diagonal elements will be strictly positive. In terms of equations, in the case of two levels, where i represents an upper level option or a nest, and j represents a lower level option, the utility function can be represented as:

$$U(i, j) = U_i + U_{ij}, \quad (5)$$

$$U(i, j) = V(i, j) + \varepsilon(i, j), \quad (6)$$

$$V(i, j) = V_i + V_{j/i}, \quad (7)$$

$$\varepsilon(i, j) = \varepsilon_i + \varepsilon_{j/i}. \quad (8)$$

It can be shown (Williams, 1977) that if $\varepsilon_{j/i}$ is Gumbel distributed, and ε_i is Logistic distributed, then the probability of choosing nest i and option j within that nest is given by Eqs. (9)–(11), where $V_{j/i}$ is the representative utility of option j within nest i considering only those attributes which differ within the nest. The scale parameters β and λ_i correspond to the upper level and nest i respectively. $C_n(i)$ represents the choice set within nest i for individual n , and consequently $C_n(0)$ represents the choice set for individual n at the upper level.

¹ To be more specific, this is a case of discrete or binary heteroskedasticity between observations, but cases of continuous heteroskedasticity might also occur.

$$P_{ij} = P_i \cdot P_{j/i} \quad (9)$$

with

$$P_{j/i} = \frac{\exp(\lambda_i V_{j/i})}{\sum_{k \in C_n(i)} \exp(\lambda_i V_{k/i})} \quad \forall i \in C_n(i), \quad (10)$$

$$P_i = \frac{\exp(\beta V_i)}{\sum_{l \in C_n(0)} \exp(\beta V_l)} \quad \forall i \in C_n(0). \quad (11)$$

The utility of options or nests at the upper level (0) is represented by (Williams, 1977):

$$V_i = X_i + \frac{1}{\lambda_i} \ln \sum_{j \in C_n(i)} \exp(\lambda_i V_{j/i}), \quad (12)$$

where X_i represents the component of utility that is associated with the attributes that are common to all options in nest i , denoted as $C_n(i)$. The structural parameter $\phi_i = \beta/\lambda_i$, is related to the correlation between options. Daganzo and Kusnic (1993) generically described this relation in analytical terms, apparently for the first time. It can be shown that ϕ_i should be in the range (0,1) for the hierarchical structure to be supported by the data. In order to make the model estimable, one of the scale parameters should be fixed. If the scale parameter of the upper level β is fixed to unity, then the scale parameter of nest i (λ_i) will be $1/\phi_i$.

A particular case of the hierarchical structure was proposed by Bradley and Daly (1997) to estimate a single model using two data sets of different variance. This arises, for example, when mixing RP and SP data within the theoretical framework developed by Ben-Akiva and Morikawa (1990a, b). The estimation includes the identification of a scale parameter (μ) equal to the ratio between the standard deviations of the two data sets; this scales all the utility values in one of the databases. Bradley and Daly (1997) demonstrated that this can be estimated as a hierarchical model where each option of one of the data sets is allocated to a single element nest and all such nests have the same scale parameter; we call this a *Single Element Nested Logit model* (SENL). Although the SENL was developed to address the problem of discrete heteroskedasticity between observations, there have also been some attempts to use it to model heteroskedasticity between options (discrete choice processes). The difference between these uses is that in the former case, each choice is made within one or other of the two homoskedastic choice sub-sets. Thus when the options associated with utility variance σ_1 are available those associated with σ_2 are not. This contrasts with the latter case in which every choice is made from a set containing options with different utility variances.

Another model has been proposed for the case of continuous heteroskedasticity between observations (Steckel and Vanhonacker, 1988). This is called Heterogeneous Conditional Logit (HCL) and, as the MNL, it is also derived assuming Extreme Value distributed error terms, but it allows individual heterogeneity by including a specific scale parameter for each individual. In this way, it provides a flexible framework to model general cases of heteroskedasticity between observations. However, in this case we would have unobserved variability between observations, and the discrete case of mixing RP and SP observations that interests us here is one of observed variability but of unknown magnitude in the population variance.

To derive a particular model, Steckel and Vanhonacker (1988) assume that the scale parameter is Gamma distributed across the population and for the special case of the exponential

distribution, establishes a closed form for the choice probabilities. Thus, it is easy to see that the HCL model is just a special case of the random coefficients or Mixed Logit (ML) family of models (Ben Akiva and Bolduc, 1996; McFadden and Train, 1997). The ML has recently been proposed as the model of the future, as apparently it is capable of approximating any discrete choice model (including the MNP) to any required degree of accuracy (Train, 1998; Ortúzar and Garrido, 1998). An application of the HCL model can be found in Bhat (1998).

As our general aim is to study the most parsimonious model which is adequate to a particular covariance structure, we will not pursue this kind of specification here, although it might be appropriate in some cases that we have not studied. Also, as the generality afforded by the ML is also provided by the Probit we have a further (admittedly post-hoc) justification for not considering the ML model further.

In order to take into account the heteroskedasticity between options problem, the *Heteroskedastic Extreme Value Logit model* (HEVL) has recently been developed and applied (Bhat, 1995; Hensher, 1996). In this model, the error terms are assumed to be mutually independent Extreme Value distributed but are allowed to have differing variances. The choice probabilities are given by:

$$P_i = \int_{w=-\infty}^{\infty} \prod_{j \in C_n, j \neq i} \Lambda \left[\frac{V_i - V_j + \theta_j w}{\theta_j} \right] \lambda(w) dw, \quad (13)$$

where $\lambda(t) = e^{-t}e^{-e^{-t}}$ and $\Lambda(t) = e^{-t}e^{-e^{-t}}$.

As in the Probit case, this integral cannot be solved analytically; however, it can be re-arranged to be evaluated using Gauss–Laguerre numerical integration (Press et al., 1992). This model has scale parameters θ_i , $i \in C_n$ which are directly related to the variance ($\sigma_i^2 = \pi^2 \theta_i^2 / 6$), one of which should be fixed in order to make the model identifiable. The IIA property does not apply to this model unless all scale parameters are equal to unity. On the contrary, Bhat (1995) established that a marginal change in the direct utility of one option would induce changes in the market shares of the others that will be smaller for those that have larger scale parameters (and variances).

3. Simulation experiments

3.1. Simulation procedure

The simulation experiments followed the general framework suggested by Williams and Ortúzar (1982), and were implemented as a realistic case where only the explanatory variables and the chosen option were available in the estimation process. The basic idea is to generate a simulated database where hypothetical individuals follow known behavioural rules. To be consistent with the theory summarised above, this was simulated assuming a population that has some fixed (and known) taste parameters and a certain error distribution. Two error distributions were used: Extreme Value in some (IID) cases and Normal with various covariance matrices in other cases.

The data sets were generated by computing the simulated choice for each observation as the option that has the largest utility (U_{in}). Those utilities were calculated as the sum of the observable

component V_{in} (i.e. the sum of the taste parameters times the corresponding attributes) plus the error term ε_{in} sampled according to the selected distribution. The attributes, generated by pseudo-random sampling, were cost and time for each of the four modes, and a binary dummy variable for high income. The time and cost attributes were normally distributed with mean and variance taken from a real database (Gaudry et al., 1989). The parameters of the utility function were also taken from models fitted to real data. Finally, the magnitude of the variances of the error terms was chosen in order to achieve a reasonable balance between the number of individuals who would change the chosen option due to the error term and those who would not. To induce heteroskedasticity, one of the variance levels was specified as twice the other.

With the data sets thus generated, considering observations consisting of the chosen option and the attribute values for the complete choice set, the choice models were estimated by maximum likelihood using GAUSS (Aptech Systems, 1994). This corresponds to the way in which models are estimated when a real database is available. We programmed a flexible Logit model, which can be used to estimate any compatible Logit structure including all those described here. The program was tested by comparing its results with those obtained from LIMDEP (Greene, 1995) and ALOGIT (Daly, 1992) for the MNL, and this gave satisfactory correspondence in all cases. We also implemented an algorithm to estimate Probit models in GAUSS using the Lerman and Manski (1981) approach with the Geweke–Hajivassilou–Keane (GHK) simulator (Hajivassiliou et al., 1993; Börsch-Supan and Hajivassiliou, 1993) for the choice probabilities, as did Bolduc (1999).

In order to investigate the consequences of adopting one or other of these models in practice, their performance was tested in terms of their ability to recover the known taste parameters, and also in terms of their prediction capabilities for scenarios representing “policy changes” which ranged from slightly different to quite different from the case used for estimation. Parameter recovery was tested by means of the well-known t -ratio test (Ortúzar and Willumsen, 1994, p. 243).

The most valuable information about the capabilities of a model is given by the response analysis, which was carried out using a range of policies. This entailed changing attribute values of the options to represent each of the policy changes and then re-executing the choice simulation procedure. In this way we generated a simulated future scenario which could be compared with the model predictions. The particular changes to the time and cost attributes for each of 17 policies P_1 to P_{17} are shown in Table 1. Policies P_1 to P_5 are rather mild, corresponding to small changes in the attribute values; on the other hand, P_6 to P_{17} represent strong policy changes, where the attribute values are changed by factors as large as two in some cases.

As mentioned above, these policies were used to evaluate the accuracy of each of the model estimates of the likely response. The error measure considered was the percentage difference between *observed* behaviour (i.e. simulated directly from the underlying choice model with the modified attribute values) and that estimated with each of the fitted models.

The minimum response error that should be considered as a prediction error is the threshold corresponding to the standard deviation of simulated *observations* generated with different seeds: this arises from the variability inherent in the simulation process. Table 2 shows the results of ten repetitions of the data generation process, each with the same attribute values and the same variance of the error term, but using different seeds for the pseudo-random number generation

Table 1

Policy changes: percentage change in the time (t_i) and cost (c_i) attribute values for option i ($1 \leq i \leq 4$)

Policy	t_1	t_2	t_3	c_1	c_2	c_3
NP						
P_1						+20%
P_2		+10%				
P_3		+30%				
P_4			+20%			
P_5						−20%
P_6						+40%
P_7		+60%				
P_8			+40%			
P_9						−40%
P_{10}		+100%				
P_{11}			+100%			
P_{12}		−50%				
P_{13}			−50%			
P_{14}						+100%
P_{15}						−50%
P_{16}		+100%	−50%		+100%	−50%
P_{17}	−50%	+100%		−50%	+100%	

Table 2

Market share observations repetitions

Standard deviation of random component Option	σ_c				$\sigma_c/2$			
	1	2	3	4	1	2	3	4
	204	355	223	218	211	450	178	161
	209	337	238	216	210	458	190	142
	216	342	245	197	201	478	176	145
	199	370	230	201	206	459	191	144
	191	370	240	199	200	461	193	146
	181	338	270	211	204	428	223	145
	205	366	229	200	202	436	193	169
	184	382	233	201	219	428	205	148
	209	362	226	203	199	435	224	142
	175	367	249	209	175	481	201	143
Mean, μ	197.3	358.9	238.3	205.5	202.7	451.4	197.4	148.5
Standard deviation, σ	13.8	15.4	13.9	7.5	11.5	19.4	16.3	9.1
Coefficient of variation, σ/μ	0.070	0.043	0.058	0.036	0.056	0.043	0.083	0.061

process. As can be seen, within each sub-sample the proportionate variability tends to decrease with increasing market share. The largest value of the proportionate variability σ/μ in this table is 0.083 for the third attribute when the error variance is $\sigma_c/2$ which arises when the market share is about 20%. Given these results and considering that when applying policy changes smaller pro-

portions are expected in some cases, we decided to take 10% as a reasonable threshold. Each discrepancy that exceeded this threshold was inspected carefully and if no other reason for it was apparent, it was considered to be an estimation error.

A goodness of fit index was also calculated to take into account the relative magnitude of the observations, as:

$$\chi^2 = \sum_i \frac{(\hat{N}_i - N_i)^2}{N_i}, \quad (14)$$

where $\hat{N}_i$ is the model estimate of the number of individuals choosing option i , and N_i is the actual (simulated) number. This index, originally proposed for this kind of use by Gunn and Bates (1982), enables us to consider the prediction error in all options. Its use here is justified by our large sample sizes which ensure that $\hat{N}_i$ is much greater than five in all cases.

3.2. Simulation results

In order to verify the workings of our procedure, we tested empirically the differences between MNL and IID Probit models, and also the differences between these kinds of models when estimated with data sets generated with Extreme Value or IID Normal error terms. These initial experiments confirmed that there are no significant differences between the MNL and independent Probit models estimated on the same data set, whether the error distribution used to generate it is IID Normal or Extreme Value. Neither are there significant differences between models of either form estimated on data generated with independent Normal and Extreme Value distributed error terms with the same variance. We note, however, that despite their identical variance–covariance specification and indistinguishable performance in policy evaluation, the IID Probit model is substantially more computationally demanding than is the MNL model, both in evaluation and in estimation.

Each of the two different sources of heteroskedasticity described above, between options and between observations, was then simulated separately. The former could arise from different levels of information about the options (e.g. typically the perception of the chosen option's attribute values is more accurate than that of seldom selected options) whilst the latter could arise when data are collected using different survey methods. Several sample sizes and covariance matrices were tested (see Munizaga, 1997); we present here some representative results.

In the case of heteroskedasticity between options, random error terms distributed IID Normal $(0, \sigma_c^2)$ were added to each option of the first pair, and error terms distributed IID Normal $(0, (\sigma_c/2)^2)$ were added to each of the others. In the second case of heteroskedasticity between observations, random error terms distributed IID Normal $(0, \sigma_c^2)$ were added to each of the 1000 members in the first group and error terms IID Normal $(0, (\sigma_c/2)^2)$ were added to each observation in the second such group. It is important to note first that the magnitude of these variances is substantial, as they imply that adding the error term would induce between 25% and 50% of the sample to change option in relation to the deterministic case. Second, that the ratio of 4.0 between these variances is also substantial, as values in practice are usually smaller (see for example Ortúzar et al., 1994; Bradley and Daly, 1997).

Heteroskedasticity between options: Some model estimation results for this case are summarised in Table 3. These results are representative of the many simulation experiments undertaken for this case, varying the number of observations, parameter values, number of repetitions of MNP estimation and the variance ratios (see Munizaga, 1997).

The target values are those used as input for the simulation which generated the database. The target value of variance σ^2 for the MNP is that of the second group of options; as only one can be estimated, we fixed the first one to the real known value (it can be easily proved that if it is fixed to unity, exactly the same results are obtained, with the variance parameter consequently scaled). The parameter θ_2 corresponds to the ratio between the standard deviations of the two sets of options.

Table 3

Model estimation results with (*t*-values against 0) and [*t*-values against target]. Heteroskedasticity between options. Database: d201, 2000 observations

	Target	MNL	SENL	MNP	HL1	HL2	HL3	HEVL
Option 1	−0.2	0.3800 (2.4)	0.2953 (0.6) [1.03]	−0.1437 (−0.7) [0.27]	−0.3516 (−1.0)	−0.0136 (−0.1)	−0.0421 (−0.1)	−0.1389 (−0.5) [0.22]
Option 2	0.6	1.1292 (15.2)	1.0469 (2.3) [0.89]	0.5081 (3.4) [−0.61]	0.6869 (3.8)	0.7505 (4.9)	0.9898 (9.4)	0.6194 (2.5) [0.08]
Option 3	0.4	0.5060 (5.0)	0.4939 (4.3) [0.58]	0.3511 (5.2) [−0.72]	0.5035 (4.7)	0.4264 (5.0)	0.5874 (5.2)	0.4258 (4.9) [0.30]
Time	−0.01	−0.0119 (−3.6)	−0.0117 (−3.3) [−0.29]	−0.0084 (−3.4) [0.65]	−0.0113 (−3.1)	−0.0104 (−3.3)	−0.0127 (−3.3)	−0.0098 (−3.2) [0.07]
Cost	−0.004	−0.0059 (−5.7)	−0.0058 (−5.4) [−1.40]	−0.0045 (−5.3) [0.59]	−0.0066 (−4.8)	−0.0055 (−5.4)	−0.0072 (−5.2)	−0.0053 (−5.4) [−1.32]
High income	1.5	1.8401 (16.2)	1.8395 (16.2) [2.10]	1.5089 (15.0) [0.09]	2.3252 (8.0)	1.8346 (16.2)	2.3778 (8.3)	1.759 (14.6) [2.11]
μ (SENL)	—	—	1.0766 (2.6)	—	—	—	—	—
σ^2 (MNP)	0.36	—	—	0.3433 (1.6) [−0.08]	—	—	—	—
ϕ_1 (HL)	—	—	—	—	1.3387 (7.0)	—	1.3742 (7.2)	—
ϕ_2 (HL)	—	—	—	—	0.5925 (3.0)	0.5102 (3.0)	—	—
θ_2 (HEVL)	0.5	—	—	—	—	—	—	0.6486 (3.9) [−0.89]
Mean log-likelihood		−1.22208	−1.22208	−1.21965	−1.22046	−1.22140	−1.22092	−1.22109

It can be seen that the MNL model yields reasonable coefficient estimates for the attributes but large errors for the option-specific constants (all the parameters are significantly different from zero according to a *t*-test at the 5% level). The SENL results are similar to those of the MNL and the scale parameter μ is close to unity, indicating that these models do not differ significantly. For this reason the SENL is not included in the response analysis discussion below.

Both the MNP and HEVL models yield a slightly better log-likelihood than the others and also better parameter estimates, especially for the option-specific constants. In the case of the SENL, MNP and HEVL models, it is possible to perform a *t*-test on the estimated parameter values for significant differences from their target values by scaling them according to the known standard deviation of the associated error term. It was found that the estimates were always consistent with the target values with the exception of the high-income dummy variable in the SENL and HEVL models, which was significantly overestimated. These comparisons were not possible for the MNL or the HL models because no scaling of their parameter values is possible that would be theoretically correct.

The HL models were included in the analysis to see whether or not our results are consistent with the empirical evidence presented by Börsch-Supan (1990) that indicates some capacity of HL models to accommodate heteroskedasticity. The fitted Hierarchical Logit models HL1 and HL3 are unacceptable on theoretical grounds because they have values greater than unity for their structural parameter ϕ_1 (see Ortúzar and Willumsen, 1994, p. 220). Because of this, these models were not considered for the response analysis. By contrast, the HL2 model is acceptable from this point of view. The results of this model fitting are consistent with the findings presented by Börsch-Supan (1990), also on the basis of synthetic studies, that heteroskedasticity rather than correlation may determine the most appropriate structure for HL models. The HL2 model fits the data significantly better than does the MNL in the present case of heteroskedasticity between options without correlation, with a Likelihood Ratio (LR) test² value for the additional parameter of 6.48 which is substantially greater than the critical χ^2 value at the 5% level of 3.84. This indicates some capability of an appropriately fitted HL model to accommodate heteroskedasticity. The HL2 model was therefore carried forward into the response analysis.

Consider now the response analysis, the full results of which are shown in Table 7 of Appendix B, with the goodness of fit index (Eq. (14)) values presented in Table 4 only for the most relevant models. In general the MNL fails to estimate accurately the response to policies affecting options three and four which have smaller variance in their error terms. The model cannot represent the greater influence that the attributes have in those options because of the smaller influence of the stochastic effect; so, their elasticity is underestimated. The χ^2 index reported for the MNL in Table 4 has large values in several cases, especially those representing strong policy measures.

² The LR test value is calculated as twice the difference between the log-likelihood at convergence of the restricted model (in this case the MNL) and that of the more general model (in this case the HL) and distributes χ^2 with a number of degrees of freedom equal to the number of linear restrictions to the more general model (Ortúzar and Willumsen, 1994, pp. 244–246).

Table 4

Summary of response analysis for heteroskedasticity between options (χ^2 index)

Policy	MNL	MNP	HL2	HEVL
NP	0.0	0.0	0.0	0.0
P_1	3.9	0.1	0.1	1.0
P_2	0.0	0.0	0.0	0.0
P_3	0.2	0.0	0.1	0.1
P_4	1.8	0.7	0.5	1.2
P_5	1.2	1.0	1.3	0.2
P_6	4.3	3.0	4.1	1.0
P_7	1.3	0.3	0.5	0.7
P_8	6.4	2.1	1.7	4.0
P_9	4.1	4.4	6.1	0.9
P_{10}	1.9	0.8	1.5	0.8
P_{11}	14.2	0.9	1.3	6.1
P_{12}	1.6	0.6	1.3	0.5
P_{13}	3.7	0.2	0.6	1.5
P_{14}	32.4	7.2	8.7	4.8
P_{15}	7.5	5.9	8.0	1.3
P_{16}	54.2	0.5	9.2	8.1
P_{17}	33.5	5.0	13.5	8.9

The critical value of χ^2 for three degrees of freedom at the 5% level is 7.815. Therefore, the MNL has significant errors at this level when modelling the effect of policies P_{11} , P_{14} , P_{16} and P_{17} which involves substantial changes in time and cost to the first two options.

The MNP gives better results, with almost all response errors within the threshold of 10% as shown in Table 7. The only cases where the MNP response error exceeds this threshold is for option three when policy changes P_6 and P_{14} are applied. However, in the case of P_6 the error is just above the limit, and in the last case the choice proportion is less than 5%, so this is a case of a relatively large proportional deviation but a small absolute one. Indeed, the χ^2 measure in Table 4 indicates a fit that is in both cases satisfactory at the 5% level of significance. The HEVL has some results that can be considered inferior to those of the MNP in terms of the percentage errors reported in Table 7. In addition, the χ^2 index in Table 4 reveals errors that are statistically significant at the 5% level in the estimation of response to policies P_{16} and P_{17} . Finally, the HL2 model yields surprisingly good results in terms of the χ^2 index. These are comparable to those of the theoretically correct MNP and HEVL models, except for the strongest policy measures P_{14} , P_{15} , P_{16} and P_{17} which are not represented well enough to have an acceptable goodness of fit χ^2 test statistic.

Heteroskedasticity between observations: The results for this case are shown in Table 5. First, two separate MNL models were estimated (and note that in each case there was a large number of observations). This corresponds to what would happen in practice if two separate data sets were used, unless there was special interest in transferring coefficients or estimating the ratio of variances. The model MNL1 was estimated with the first group of 1000 observations, for which $\sigma = \sigma_c$, whilst the MNL2 model was estimated with the second such group for which $\sigma = \sigma_c/2$.

Table 5

Model estimation results with (t -values against 0) and [t -values against target]. Heteroskedasticity between observations. Database: d204, 2000 observations

Parameter	Target	MNL1	MNL2	MNL	SENL	MNP
Option 1	−0.2	−0.3001 (−1.3) [0.37]	−0.4294 (−1.5) [−0.01]	−0.3110 (−1.8)	−0.2362 (−1.9) [−0.18]	−0.1418 (−1.4) [0.57]
Option 2	0.6	0.5913 (6.0) [−0.51]	1.2535 (12.0) [−0.28]	0.9121 (2.9)	0.6269 (9.8) [−0.22]	0.5798 (0.3) [−0.01]
Option 3	0.4	0.3574 (2.7) [−0.53]	0.8175 (5.4) [−0.25]	0.5560 (5.6)	0.4313 (5.7) [0.05]	0.3557 (6.0) [−0.75]
Time	−0.01	−0.0101 (−2.1) [−0.12]	−0.0221 (−4.5) [0.15]	−0.0156 (−4.6)	−0.0110 (−4.7) [−0.13]	−0.0105 (−5.1) [−0.24]
Cost	−0.004	−0.0048 (−3.2) [0.35]	−0.0106 (−6.1) [1.18]	−0.0072 (−6.4)	−0.0053 (−6.4) [−1.24]	−0.0046 (−6.5) [−0.85]
High income	1.5	1.8552 (10.9) [1.48]	3.5053 (15.4) [1.31]	2.5545 (19.3)	1.7910 (12.3) [1.29]	1.5252 (12.6) [0.21]
μ (SENL)	2.0	—	—	—	1.9709 (11.0) [−0.26]	—
σ^2 (MNP)	0.36	—	—	—	—	0.3690 (5.7) [0.14]
Mean log-likelihood		−1.28920	−1.08936	−1.20585	−1.18945	−1.18794

The parameters of these models should be divided by $\lambda = \pi/\sqrt{6}\sigma$ in order to make comparisons with the target values. In the case of model MNL1 with $\sigma = 1.2$ this gives $\lambda = 1.0688$ whilst for model MNL2 with $\sigma = 0.6$ this gives $\lambda = 2.1376$. The theoretically correct (for this case) SENL model was also estimated, allowing the estimation of an extra parameter (μ) representing the ratio between the standard deviations of the error terms in the two data sets. The same was done for the MNP model, where an error structure was defined in a way such that some observations had a fixed value for the variance (to represent the scaling) and the reminder had an unknown value (to be estimated).

The results for MNL1 and MNL2 show that both the likelihood and the parameter estimates are better for MNL2 and this is due to the smaller variance of the sample generation process. The estimation of a single MNL model for both data sets together yields parameter estimates that have intermediate values between those of MNL1 and MNL2. No scale parameter is available for this model, so comparisons cannot be made against the target values. However, the response analysis shows that this model is remarkably robust and that the estimation errors are similar to those of the separate models (see Tables 8 and 9 and also the summary χ^2 index values in Table 6).

Table 6

Summary of response analysis in some cases for heteroskedasticity between observations (χ^2 index)

Policy	MNL	SENL	MNP
NP	0.0	0.1	0.1
P_1	0.2	0.4	0.2
P_2	0.2	0.3	0.2
P_3	1.0	1.0	0.7
P_4	0.5	0.5	0.4
P_5	0.3	0.1	0.4
P_6	0.4	0.9	0.9
P_7	2.5	2.3	1.4
P_8	0.9	1.1	0.4
P_9	0.2	0.2	0.3
P_{10}	0.7	0.6	0.1
P_{11}	0.3	0.7	0.2
P_{12}	0.2	0.7	1.0
P_{13}	1.3	1.2	0.6
P_{14}	0.3	0.6	1.1
P_{15}	0.1	1.2	0.7
P_{16}	2.4	6.2	2.4
P_{17}	3.1	1.9	7.4

The SENL coefficients scaled by $\lambda = 1.0688$ are similar to those estimated by MNL1 and MNL2, and the variance ratio (μ) is recovered accurately. The likelihood of this model is better than that of the MNL but despite all of this, the response analysis indicates no significant advantage in terms of predictive capabilities. The MNP also has better likelihood than the MNL and recovered the parameter values accurately; none of the estimates differed significantly from their target values. Despite this, the MNP model was no better in response analysis than the MNL. Indeed, although none of the χ^2 summary indices were significantly different from zero at the 5% level, the MNP had the largest value (7.4) in the case of policy P_{17} .

It is important to mention that we rejected the hypothesis of there being an influence of the relative sample sizes on the response analysis. This was investigated by performing the same experiments with databases of different sizes and compositions (see Munizaga, 1997). The parameter estimates were different because of the variations in scaling that resulted from different proportions of observations with small and large variances, but the response analyses led to broadly the same conclusions.

4. Discussion and conclusions

Several forms of Logit models and a Multinomial Probit (MNP) model estimated using the GHK simulator have been compared in terms of (a) their recovery of the parameters used to synthesise the data, and (b) their capabilities to estimate response to policy changes. This was done for various cases including ones in which the assumptions of the MNL model of independent and identically distributed error terms did not hold.

Two different kinds of heteroskedasticity were artificially incorporated into the data generation process: heteroskedasticity between options and heteroskedasticity between observations. For the former kind, both the MNP and the HEVL models are theoretically correct, whilst for the latter both the SENL and the MNP models are consistent with the error structure.

The results of these tests show that discrete heteroskedasticity between observations can be accommodated in practice by the MNL model even though in theory it is not adequate for such cases. The response analysis shows that this model can estimate the effect of large policy changes with good accuracy. In this case, the advantage of using SENL or MNP would be in respect of their capability to quantify the ratio between the error variances associated with the different sources of data.

We note, however, that in the example investigated here we introduced only heteroskedasticity so that the coefficients of the attributes were identical between the samples, which may be unusual in practice. It is common to find that the results of MNL and SENL models applied to real data sets differ substantially, even yielding corresponding coefficients that have opposite signs (see for example Ortúzar et al., 1994). This of course implies different elasticities. We conclude from this that those differences are not due to heteroskedasticity between observations alone and hence that it is necessary to do further research using real data to understand and identify their main source.

In the case of heteroskedasticity between options, which can be found for instance in ranking SP experiments, the MNL model gives less accurate estimates of the effects of policy changes. Although it gives fairly good estimates of coefficients of attributes, it does not recover accurately the option-specific constants. Only policy changes affecting the utility of options that have larger error variance were evaluated with acceptable accuracy in this case. The SENL model performed no better with this kind of heteroskedasticity.

The HL model, which Börsch-Supan (1990) reported as having some capability to accommodate heteroskedasticity between options, showed surprisingly good behaviour in terms of response analysis, with response errors that were as good as those of the best models for most mild policy measures and were substantially worse only for the strongest ones. As expected, the theoretically correct models, which are the MNP and the HEVL for this case, showed better results than the others. Both of these models recovered accurately the target parameter values and performed satisfactorily in terms of the response analysis, though the HEVL could not estimate response faultlessly to our strongest test policies.

The SENL model fitted the data with heteroskedasticity between observations better than did the MNL model, with more accurate recovery of model parameters and better likelihood values. Despite this, the estimates of response to policy variations were no better. Furthermore, where there is heteroskedasticity between options, this model was found to offer no advantage over the MNL model. The MNP model fitted the data with heteroskedasticity between observations slightly better than did the SENL model, yielding good estimates of model parameters and slightly (though significantly) better likelihood. However, the MNP still had no advantage over the MNL model in estimating the response to policy variations in these circumstances. Rather, the main strength of the MNP model lies in the case of heteroskedasticity between options, in which case this was the only model that recovered all parameter values reliably and provided good estimates of response to policy variations for all the tests applied.

We conclude that the MNL model is remarkably robust and can be used reliably to evaluate the effects of even substantial policy changes in the presence of heteroskedasticity when this lies between observations. In this case its performance in response analysis is comparable to the more computationally demanding MNP model. Although the HL model performed surprisingly well in the case of heteroskedasticity between options, in these experiments, the SENL and HL models were found to offer no substantial advantages over the MNL model. These results give further credentials to the Logit family (see Williams and Ortúzar, 1982).

The MNP and HEVL models can both accommodate heteroskedasticity between options in estimating utility functions and response to policy variations. However, as the HEVL model might be expected to inherit some of the advantages of the MNL (due to its Logit genesis), it is worth mentioning that its estimation requires a similar computational effort (in our implementation) to that of the corresponding independent MNP model that has a diagonal covariance matrix. On the basis of this, we find that use of the MNP model is justified in these circumstances, as its use is now a practical proposition using appropriate simulation techniques and contemporary computers.

Acknowledgements

We are grateful for the financial support received from FONDECYT, CONICYT, The British Council, Fundación Andes and the UK Engineering and Physical Science Research Council. The hospitality of the University of London Centre for Transport Studies during part of this research is gratefully acknowledged by the two Chilean authors. We thank Moshe Ben-Akiva, Mark Bradley, Andrew Daly and Taka Morikawa, who were kind enough to advise us by private communication. We also thank Vassilis Hajivassiliou, who made available several publications and the GAUSS code to evaluate the multivariate Normal distribution by anonymous ftp. Finally we wish to thank two anonymous referees for their incisive and helpful comments.

Appendix A. Estimation of Probit models

Estimation of Multinomial Probit models (MNP) is now possible in a reasonable time using current computational techniques due to some recent advances in related fields. The main difficulty in estimation is due to the lack of an analytical expression for the probability function of a variable which is Multivariate Normal (MVN) distributed. In general, this leads to a multivariate integral that cannot be solved analytically so that either numerical approximation or simulation must be used to evaluate it.

One of the approaches proposed for the problem of estimating probabilities for the MVN distribution is the use of Clark's approximation (Daganzo et al., 1977). This has the advantage of being a simple analytical formula, but unfortunately it suffers from severe bias where there is heteroskedasticity, random taste variation or small choice proportions (Horowitz et al., 1984). Lerman and Manski (1981) proposed the method of Simulated Maximum Likelihood (SML) where at each iteration the likelihood is replaced directly by a simulated probability. Alter-

natively, McFadden (1989), and Pakes and Pollard (1989) proposed, at the same time, the Method of Simulated Moments (MSM). This avoids the multiple integral by replacing the choice probability in the moments equation by an unbiased value that can be estimated by simulation.

Lerman and Manski (1981) used the classic frequency simulator, but that method has two major limitations. First, given a finite number of repetitions, there remains a non-zero probability of recording a zero frequency. This implies that the estimators are consistent only in the limit that both the sample size and the number of repetitions tend to infinity. The second limitation is that the simulator is not continuous so that a change in the parameters can induce a discrete change in the frequency. Thus to obtain good estimates of small probabilities requires that a large number of repetitions be made. However, we have recently learnt about another use of the frequency simulator, implemented by Lam and Mahmassani (1997). These authors have concentrated on the optimisation problem associated to the maximum likelihood search, developing an efficient code that allows tackling problems with an unlimited number of parameters and options, including any form of utility function. This has been further enhanced by the use of parallel computing, with good results in terms of efficiency. A recent application to a heteroskedastic problem with 42 options and 39 parameters can be found in Garrido and Mahmassani (1998).

One of the most popular alternative simulators is due to Stern (1992) and is based on separating the error term into two components. One component is normally distributed with zero mean and diagonal covariance matrix, and the other has a covariance matrix that is as small as possible. The method has a precision (in terms of the variance of the simulated probabilities) that depends on the relative magnitude of the two components of the error term. In the limit that the second component is zero, this method is exact.

Using advances in integration techniques based on Monte Carlo simulation, Börsch-Supan and Hajivassiliou (1993) proposed the Geweke–Hajivassilou–Keane (GHK) simulator that produces unbiased, continuous and differentiable probabilities. The values lie strictly within the (0,1) interval, and the computing time required increases only linearly with the number of options and is independent of the true probabilities. This simulator, based on a recursive reduction of the problem dimension, requires the generation of repetitions of a unidimensional truncated Normal distribution. Börsch-Supan and Hajivassiliou (1993) implemented the SML method and showed that the variance obtained from the GHK simulator is smaller than that of either the frequency simulator or Stern's simulator. For these reasons, we adopt the GHK simulator for the numerical work presented in this paper. An implementation of the GHK simulator to estimate Probit models by maximum simulated likelihood has been reported in Bolduc (1999), where the details of the estimation procedure are discussed and the model is estimated for the Santiago database (Gaudry et al., 1989).

Appendix B. Full results of simulation processes

See Tables 7–9 (overleaf) for full results of simulation processes.

Table 7
Heteroskedasticity between options. Estimated choice numbers and percentage difference between them and target values

Policy	Target value				MNL				SEN				Probit				HL2				HEVL			
	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
NP	497	849	344	310	495.5	850.3	343.7	310.4	497.2	848.8	344.1	309.9	496.1	848.7	345.6	309.5	496.7	849.2	343.7	310.4	497.7	846.6	344.3	310.3
P_1	516	872	266	346	-0.3	0.2	-0.1	0.1	0.3	-0.2	0.1	-0.2	-0.2	-0.0	0.5	-0.2	-0.1	0.0	-0.1	0.1	0.4	-0.4	0.2	-0.0
					509.4	880.3	288.8	321.4	512.0	880.3	286.2	321.5	511.4	880.0	265.1	343.4	509.2	875.8	265.5	349.5	513.1	879.6	275.1	331.0
P_2	502	830	351	317	-1.3	1.0	8.6	-7.1	-0.8	1.0	7.6	-7.1	-0.9	0.9	-0.3	-0.8	-1.3	0.4	-0.2	1.0	0.7	-0.1	-4.7	3.0
					503.8	828.2	351.0	317.0	505.4	827.0	351.2	329.1	500.1	830.9	351.9	317.0	504.1	829.8	350.0	316.1	503.9	826.2	351.7	317.1
P_3	514	793	366	327	0.4	-0.2	-	-	0.7	-0.4	0.1	3.8	-0.4	0.1	0.3	-	0.4	-0.0	-0.3	-0.3	0.0	-0.2	0.2	0.0
					520.1	784.6	365.3	330.0	521.5	784.1	365.2	329.1	511.2	793.2	365.8	329.8	518.4	791.5	362.6	327.5	516.0	786.1	366.4	330.5
P_4	506	864	302	328	1.2	-1.1	-0.2	0.9	1.5	-1.1	-0.2	0.6	-0.5	0.0	-0.1	0.9	0.9	-0.2	-0.9	0.2	-0.8	0.2	0.3	0.2
					501.2	862.5	321.4	314.9	503.2	861.5	320.4	314.7	501.0	861.5	314.6	322.9	501.6	859.6	313.6	325.2	503.5	859.1	318.2	318.1
P_5	477	818	421	284	-0.9	-0.2	6.4	-4.0	-0.6	-0.3	6.1	-4.1	-1.0	-0.3	4.2	-1.6	-0.9	-0.5	3.8	-0.9	0.5	-0.4	-1.0	1.0
					479.3	816.5	406.7	297.8	479.9	812.5	411.0	296.6	474.2	813.6	437.5	274.7	481.1	816.2	434.4	268.3	479.2	807.8	425.2	286.8
P_6	527	888	224	361	0.5	-0.2	-3.4	4.9	0.6	-0.7	-2.4	4.4	-0.6	-0.5	3.9	-3.3	0.9	-0.2	3.2	-5.5	-0.0	-1.1	4.5	-3.7
					521.4	906.2	241.4	331.0	524.5	907.3	236.7	331.5	526.5	902.7	200.1	370.6	518.9	896.6	200.6	384.0	525.7	907.1	217.2	348.7
P_7	530	742	390	338	-1.1	2.0	7.8	-8.3	-0.5	2.2	5.7	-8.2	-0.1	1.7	-10.7	2.7	-1.5	1.0	-10.4	6.4	0.8	0.1	-10.0	5.4
					543.7	720.9	386.3	349.1	544.7	721.6	385.9	347.8	528.7	736.7	387.1	347.6	539.3	735.4	381.0	344.3	533.3	727.9	387.9	350.0
P_8	517	873	266	344	2.6	-2.8	-0.9	3.3	2.8	-2.8	-1.1	2.9	-0.2	-0.7	-0.7	2.8	1.8	-0.9	-2.3	1.9	-1.9	1.0	0.4	0.3
					506.6	874.0	300.3	319.1	508.9	873.8	298.2	319.1	508.6	871.0	286.7	333.6	506.1	869.1	285.3	339.5	509.0	870.8	293.6	325.4
P_9	454	783	505	258	-2.0	0.1	12.9	-7.2	-1.6	0.1	12.1	-7.2	-1.6	-0.2	7.8	-3.0	-2.1	-0.4	7.3	-1.3	0.5	-0.4	-2.2	2.0
					460.8	777.6	477.9	283.7	460.1	771.3	487.1	281.5	452.4	770.7	540.2	236.7	462.2	776.9	535.3	225.6	457.6	763.2	517.4	260.9
P_{10}	554	657	427	362	1.5	-0.7	-5.4	9.9	1.3	-1.5	-3.5	9.1	-0.4	-1.6	7.0	-8.3	1.8	-0.8	6.0	-12.6	-0.7	-1.8	8.3	-8.0
					573.2	640.4	413.1	373.4	573.8	642.2	412.2	371.7	548.7	665.1	413.9	372.3	565.7	663.7	404.7	365.9	554.7	654.2	415.3	375.0
P_{11}	531	899	200	370	3.5	-2.5	-3.3	3.1	3.6	-2.3	-3.5	2.7	-1.0	1.2	-3.1	2.8	2.1	1.0	-5.2	1.1	-3.2	2.2	0.5	0.4
					520.8	904.7	244.1	330.4	523.8	905.8	239.5	330.9	524.1	898.3	212.0	365.5	517.2	892.9	211.7	378.2	523.1	901.4	229.2	345.0
P_{12}	473	951	310	266	-1.9	0.6	22.1	-10.7	-1.4	0.8	19.8	-10.6	-1.3	-0.1	6.0	-1.2	-2.6	-0.7	5.8	2.2	0.4	-0.4	-6.1	4.4
					452.4	963.1	307.2	277.3	454.8	959.6	308.2	277.4	467.0	944.7	310.7	277.7	459.0	947.7	311.9	281.4	465.2	951.1	306.5	275.9
P_{13}	473	813	436	278	-4.4	1.3	-0.9	4.2	-3.8	0.9	-0.6	4.3	-1.3	-0.7	0.2	4.4	-3.0	-0.3	0.6	5.8	2.8	-1.2	-0.2	-0.5
					479.7	816.9	405.4	298.1	480.3	813.3	409.5	296.9	477.8	816.1	428.9	277.2	482.5	819.2	426.6	271.7	481.2	812.0	416.5	289.3
P_{14}	551	990	95	424	1.4	0.5	-7.0	7.2	1.5	0.0	-6.1	6.8	1.0	0.4	-1.6	-0.3	2.0	0.8	-2.2	-2.3	0.3	-0.6	2.7	-3.0
					547.3	963.0	137.9	351.9	551.0	965.5	130.6	352.9	551.2	946.3	72.7	429.7	535.6	932.8	78.2	453.4	550.2	961.6	101.1	385.0
P_{15}	445	762	552	241	-0.7	-2.7	45.2	-17.0	-	-2.5	37.5	-16.8	0.0	-4.4	-23.5	1.3	-2.8	-5.8	-17.7	6.9	0.5	-0.1	-26.7	9.4
					450.7	756.7	516.6	276.0	449.2	749.0	528.6	273.3	439.5	747.5	595.2	217.9	451.5	755.0	588.7	204.8	445.7	738.9	567.3	247.3
P_{16}	496	317	928	259	1.3	-0.7	-6.4	14.5	0.9	-1.7	-4.2	13.4	-1.2	-1.9	7.8	-9.6	1.5	-0.9	6.6	-15.0	-1.1	-2.4	9.8	-10.4
					545.5	299.8	804.9	350	537.6	298.5	822.1	341.9	494.6	305.8	937.9	261.6	538.6	325.6	912.9	222.9	504.7	308.1	888.4	298.4
P_{17}	900	371	387	342	10.0	-5.4	-13.3	35.1	8.4	-5.8	-11.4	32.0	-0.3	-3.5	1.1	1.0	8.6	2.7	-1.6	-13.9	-7.5	2.8	10.4	-14.7
					1009.0	286.3	369.9	334.3	1003.2	291.7	370.8	334.3	933.9	334.2	385	346.1	971.4	319.1	372.7	336.9	950.1	324.2	380.7	346.7
					12.1	-22.8	-4.4	-2.3	11.5	-21.4	-4.2	-2.3	3.8	-9.9	-0.5	1.2	7.9	-14.0	-3.7	-1.5	-5.8	13.2	2.9	3.7

Table 8

Response analysis for separate models. Database d204 1000–1000 obs. Percentage difference between estimated and target values

Policy	Target value				MNL1 ($\sigma = 1.2$)				Target value				MNL2 ($\sigma = 0.6$)			
	Model1	Model2	Model3	Model4	Model1	Model2	Model3	Model4	Model1	Model2	Model3	Model4	Model1	Model2	Model3	Model4
NP	204	355	223	218	203.8	355.2	223.0	218.0	211	450	178	161	211.5	449.5	178.2	160.8
					−0.1	0.1	—	—					0.2	−0.1	0.1	−0.1
P_1	216	371	182	231	210.1	368.6	195.0	226.3	217	471	139	173	219.0	478.2	131.4	171.4
					−2.7	−0.6	7.1	−2.0					0.9	1.5	−5.5	−0.9
P_2	206	340	230	224	206.3	346.3	226.2	221.2	216	429	188	167	216.4	429.9	185.8	167.9
					0.1	1.9	−1.7	−1.3					0.2	0.2	−1.2	0.5
P_3	212	313	241	234	211.1	328.8	232.6	227.5	224	391	206	179	225.8	391.5	200.9	181.7
					−0.4	5.0	−3.5	−2.8					0.8	0.1	−2.5	1.5
P_4	211	365	197	227	206.5	360.8	211.3	221.5	213	459	161	167	214.7	461.9	158.0	165.4
					−2.1	−1.2	7.3	−2.4					0.8	0.6	−1.9	−1.0
P_5	195	337	260	208	196.8	340.5	253.7	209.0	198	415	236	151	201.6	413.7	237.0	147.7
					0.9	1.0	−2.4	0.5					1.8	−0.3	0.4	−2.2
P_6	220	384	158	238	215.8	380.7	169.8	233.7	221	487	115	177	224.5	500.4	95.4	179.7
					−1.9	0.9	7.5	−1.8					1.6	2.8	−17.0	1.5
P_7	221	281	253	245	218.0	303.5	241.9	236.6	239	333	232	196	238.6	336.8	222.7	201.8
					−1.4	8.0	−4.4	−3.4					−0.2	1.1	−4.0	3.0
P_8	217	374	179	230	209.0	366.2	200.0	224.8	215	469	144	172	217.6	473.1	139.7	169.6
					−3.7	−2.1	11.7	−2.3					1.2	0.9	−3.0	−1.4
P_9	188	321	293	198	189.1	324.6	287.1	199.2	183	384	299	134	189.0	371.1	307.8	132.1
					0.6	1.1	−2.0	0.6					3.3	−3.4	2.9	−1.4
P_{10}	230	249	261	260	226.6	271.8	253.6	248.1	254	283	251	212	253.0	271.3	249.3	226.4
					−1.5	9.2	−2.8	−4.6					−0.4	−4.1	−0.7	6.8
P_{11}	222	388	150	240	216.0	380.9	169.2	233.9	224	494	102	180	224.4	500.2	95.8	179.6
					−2.7	−1.8	12.8	−2.5					0.2	1.3	−6.1	−0.2
P_{12}	194	404	201	201	190.9	401.4	206.2	201.5	185	536	147	132	184.4	548.8	140.5	126.4
					−1.6	−0.6	2.6	0.2					−0.3	2.4	−4.4	−4.2
P_{13}	193	333	269	205	196.6	340.3	254.3	208.8	195	410	245	150	201.6	413.0	238.0	147.4
					1.9	2.2	−5.5	1.9					3.4	0.7	−2.9	−1.7
P_{14}	234	412	102	252	229.1	409.6	109.8	251.6	235	528	36	201	233.6	538.1	34.6	193.7
					−2.1	−0.6	7.6	−0.2					−0.6	1.9	−3.9	−3.6
P_{15}	186	313	308	193	185.0	316.2	304.7	194.0	180	365	327	128	181.7	347.6	347.2	125.6
					−0.538	1.0	−1.1	0.5					0.9	−4.8	6.2	−1.9

Table 9

Heteroskedasticity between observations. 1000/1000 obs. $\sigma_1 = 1.2$, $\sigma_2 = 0.6$. Percentage difference between estimated and target values

Policy	Target value				Multinomial logit				SENL				Probit			
	M1	M2	M3	M4	Model1	Model2	Mode3	Mode4	Model1	Model2	Mode3	Mode4	Model1	Model2	Mode3	Mode4
NP	415	805	401	379	415.8	804.3	401.4	378.5	410.8	809.8	400.6	378.8	413.5	808.9	402.1	375.8
					0.2	−0.1	0.1	−0.1	−1.0	0.6	−0.1	−0.0	−0.4	0.5	0.3	−0.8
P_1	433	842	321	404	430.4	844.7	327.1	397.7	425.0	852.8	323.9	398.3	429.1	849.0	324.2	397.9
					−0.6	0.3	1.9	−1.5	−1.9	1.3	0.9	−1.4	−0.9	0.8	1.0	−1.5
P_2	422	769	418	391	423.4	776.2	411.9	388.5	418.3	780.8	411.7	389.3	421.5	778.9	413.3	386.5
					0.3	0.9	−1.5	−0.6	−0.9	1.5	−1.5	−0.4	−0.1	1.3	−1.1	−1.2
P_3	436	704	447	413	438.1	721.2	432.5	408.1	432.7	723.9	433.5	409.9	437.1	720.0	435.6	407.6
					0.5	2.4	−3.2	−1.2	−0.7	2.8	−3.0	−0.8	0.3	2.3	−2.6	−1.3
P_4	424	824	358	394	422.2	822.0	368.8	386.9	416.8	828.0	368.1	387.1	420.7	827.6	365.9	386.0
					−0.4	−0.2	3.0	−1.8	−1.7	0.5	2.8	−1.8	−0.8	0.4	2.2	−2.0
P_5	393	752	496	359	398.4	757.7	487.6	356.3	393.5	758.2	492.4	355.9	395.2	762.1	492.1	350.9
					1.4	0.8	−1.7	−0.7	0.1	0.8	−0.7	−0.9	0.6	1.3	−0.8	−2.3
P_6	441	871	273	415	442.6	879.0	264.3	414.1	436.5	888.0	261.0	414.5	442.1	882.5	258.7	416.9
					0.4	0.9	−3.2	−0.2	−1.0	2.0	−4.4	−0.1	0.2	1.3	−5.2	0.5
P_7	460	614	485	441	458.7	642.6	462.2	436.4	452.7	642.6	464.9	439.7	458.9	635.3	467.8	438.3
					−0.3	4.7	−4.7	−1.0	−1.6	4.7	−4.1	−0.3	−0.2	3.5	−3.5	−0.6
P_8	432	843	323	402	428.2	838.5	338.5	394.8	422.3	844.9	338.1	394.7	427.5	844.9	332.3	395.6
					−0.9	−0.5	4.8	−1.8	−2.2	0.2	4.7	−1.8	−1.0	0.2	2.9	−1.6
P_9	371	705	592	332	378.2	705.0	585.5	331.3	372.7	698.5	599.2	329.6	374.2	709.2	593.3	323.7
					1.9	0.0	−1.1	−0.2	0.5	−0.9	1.2	−0.7	0.9	0.6	0.2	−2.5
P_{10}	484	532	512	472	483.3	546.3	498.9	471.5	476.1	543.6	503.7	476.6	484.9	531.1	507.7	476.5
					−0.1	2.7	−2.6	−0.1	−1.6	2.2	−1.6	1.0	0.2	−0.2	−0.8	1.0
P_{11}	446	882	252	420	443.5	881.3	260.0	415.2	436.4	887.5	261.8	414.3	444.5	888.5	246.7	420.5
					−0.6	−0.1	3.2	−1.1	−2.2	0.6	3.9	−1.4	−0.3	0.7	−2.1	0.1
P_{12}	379	940	348	333	375.2	949.5	347.8	327.6	370.2	957.9	345.3	326.6	371.0	961.6	345.4	322.3
					−1.0	1.0	−0.1	−1.6	−2.3	1.9	−0.8	−1.9	−2.1	2.3	−0.7	−3.2
P_{13}	388	743	514	355	397.4	755.1	492.4	355.1	393.4	757.7	493.3	355.6	392.9	756.0	503.7	347.7
					2.4	1.6	−4.2	0.0	1.4	2.0	−4.0	0.2	1.3	1.7	−2.0	−2.1
P_{14}	469	940	138	453	467.2	950.2	134.6	448.1	459.3	956.4	137.1	447.2	468.8	948.6	126.1	456.7
					−0.4	1.1	−2.5	−1.1	−2.1	1.7	−0.7	−1.3	−0.0	0.9	−8.6	0.8
P_{15}	366	678	635	321	367.0	676.6	638.6	317.8	361.0	666.1	657.5	315.4	362.7	680.8	647.5	309.4
					0.3	−0.2	0.6	−1.0	−1.4	−1.8	3.5	−1.7	−0.9	0.4	2.0	−3.6
P_{16}	421	182	989	408	419.3	191.6	1006.5	382.6	406.4	183.7	1038.1	371.8	417.8	176.0	1020.5	385.7
					−0.4	5.3	1.8	−6.2	−3.5	0.9	5.0	−8.9	−0.8	−3.3	3.2	−5.5
P_{17}	851	227	478	444	889.6	216.8	459.5	434.1	875.6	211.4	469.1	444.0	906.0	201.4	460.2	432.4
					4.5	−4.5	−3.9	−2.2	2.9	−6.9	−1.9	—	6.5	−11.3	−3.7	−2.6

References

- Aptech Systems, 1994. GAUSS User's Manuals. Maple Valley.
- Ben-Akiva, M.E., Bolduc, D., 1996. Multinomial Probit with a Logit kernel and a general parametric specification of the covariance structure. Working Paper, Department d'Economique, Université Laval, Quebec.
- Ben-Akiva, M.E., Lerman, S.R., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge, MA.

- Ben-Akiva, M.E., Morikawa, T., 1990a. Estimation of travel demand models from multiple data sources. In: Koshi, M. (Ed.), *Transportation and Traffic Theory*. Elsevier, New York.
- Ben-Akiva, M.E., Morikawa, T., 1990b. Estimation of switching models from revealed preferences and stated intentions. *Transportation Research* 24A (6), 485–495.
- Bhat, C.R., 1995. A heteroskedastic extreme value model of intercity travel mode choice. *Transportation Research* 29B (6), 471–483.
- Bhat, C.R., 1998. Accommodating flexible substitution patterns in multi-dimensional choice modelling: formulation and application to travel mode and departure time choice. *Transportation Research* 32B (7), 455–466.
- Bolduc, D., 1992. Generalised autoregressive errors in the multinomial Probit model. *Transportation Research* 26B (2), 155–170.
- Bolduc, D., 1999. A practical technique to estimate multinomial Probit models in transportation. *Transportation Research* 33B (1), 63–79.
- Börsch-Supan, A., 1990. Recent developments in flexible discrete choice models: nested Logit analysis versus simulated moments Probit analysis. In: Fisher, M.M., Nijkamp, P., Papageorgiou, Y.Y. (Eds.), *Behavioural Modelling of Spatial Choices and Processes*. North-Holland, Amsterdam.
- Börsch-Supan, A., Hajivassiliou, V.A., 1993. Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. *Journal of Econometrics* 58 (3), 347–368.
- Bradley, M.A., Daly, A.J., 1997. Estimation of Logit choice models using mixed stated preference and revealed preference information. In: Stopher, P., Lee-Gosselin, M. (Eds.), *Understanding Travel Behaviour in an Era of Change*. Pergamon Press, Oxford.
- Daganzo, C.F., Bouthelier, F., Sheffi, Y., 1977. Multinomial Probit and qualitative choice: a computationally efficient algorithm. *Transportation Science* 11 (4), 338–358.
- Daganzo, C.F., Kusnic, M., 1993. Two properties of the nested Logit model. *Transportation Science* 27 (4), 395–400.
- Daly, A.J., 1987. Estimating “tree” Logit models. *Transportation Research* 21B (4), 251–267.
- Daly, A.J., 1992. ALOGIT 3.2. User’s Guide. Hague Consulting Group, The Hague.
- Garrido, R.A., Mahmassani, H., 1999. Forecasting freight transportation demand with the space–time Multinomial Probit model. *Transportation Research B* (in press).
- Gaudry, M.J.I., Jara-Díaz, S.R., Ortúzar, J de D., 1989. Value of time sensitivity to model specification. *Transportation Research* 23B (2), 151–158.
- Greene, W.H., 1995. *LIMDEP Version 7.0 User’s Manual*. Econometric Software, Bellport.
- Gumbel, E.J., 1957. *Statistics of Extremes*. Columbia University Press, New York.
- Gunn, H.F., Bates, J.J., 1982. Statistical aspects of travel demand modelling. *Transportation Research* 16A (5/6), 371–382.
- Hajivassiliou, V.A., McFadden, D., Ruud, P., 1993. Simulation of multivariate normal rectangle probabilities and their derivatives: theoretical and computational results. Working Paper. Cowles Foundation, Yale University.
- Hensher, D., 1996. Extending valuation to controlled value functions and non-uniform scaling with generalised unobserved variances. Working Paper ITS-WP-96-9, Institute of Transport Studies, The University of Sydney.
- Hensher, D.A., Johnson, L.W., 1981. *Applied Discrete Choice Modelling*. Croom Helm, London.
- Horowitz, J.L., Sparmann, J.M., Daganzo, C.F., 1984. An investigation of the accuracy of the Clark approximation for the Multinomial Probit model. *Transportation Science* 16 (3), 383–401.
- Lam, S.H., Mahmassani, H., 1997. Multinomial Probit model estimation: computational procedures and applications. In: Stopher, P., Lee-Gosselin, M. (Eds.), *Understanding Travel Behaviour in an Era of Change*. Pergamon Press, Oxford.
- Lerman, S., Manski, C., 1981. On the use of simulated frequencies to approximate choice probabilities. In: Manski, C., McFadden, D. (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge, MA.
- McFadden, D., 1974. Conditional Logit analysis of qualitative choice behaviour. In: Zarembka, P. (Ed.), *Frontiers in Econometrics*. Academic Press, New York.
- McFadden, D., 1989. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* 57 (5), 995–1026.
- McFadden, D., Train, T., 1997. Mixed MNL models for discrete response. Working Paper, Department of Economics, University of California at Berkeley.

- Munizaga, M., 1997. Implicancias de la Naturaleza de los Datos en la Modelación de Elecciones Discretas. Ph.D. thesis, Department of Transport Engineering, Pontificia Universidad Católica de Chile (in Spanish).
- Ortúzar, J.de D., Garrido, R.A., 1998. Methodological developments. Workshop Report, Eighth International Conference on Travel Behaviour, Austin, Texas.
- Ortúzar, J.de D., Willumsen, L.G., 1994. *Modelling Transport*, 2nd ed. Wiley, Chichester.
- Pakes, A., Pollard, D., 1989. Simulation and the asymptotics of optimisation estimators. *Econometrica* 57 (5), 1027–1057.
- Pearmain, D., Swanson, J., Kroes, E., Bradley, M., 1991. *Stated Preference Techniques: A Guide to Practice*. Steer Davies Gleave and Hague Consulting Group, London.
- Press, W.H., Teukolski, S.A., Vetterling, W.T., Flannery, B.P., 1992. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 2nd ed. Cambridge University Press, Cambridge.
- Steckel, J.H., Vanhonacker, W.R., 1988. A heterogeneous conditional Logit model of choice. *Journal of Business and Economic Statistics* 6 (3), 391–398.
- Stern, S., 1992. A method for smoothing simulated moments of discrete probabilities in Multinomial Probit models. *Econometrica* 60 (5), 943–952.
- Train, K.E., 1998. Recreation demand models with taste differences over people. *Land Economics* 74 (2), 230–239.
- Williams, H.C.W.L., 1977. On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning 9A* (3), 285–344.
- Williams, H.C.W.L., Ortúzar, J.de D., 1982. Behavioural theories of dispersion and the mis-specification of travel demand models. *Transportation Research 16B* (3), 167–219.