Fac. CC. Económicas y Empresariales
Universidad de La Laguna

Fac. CC. Económicas y Empresariales
Univ. de Las Palmas de Gran Canaria

Preference heterogeneity and willingness to pay for travel time

Francisco Javier Amador Morera*
Rosa Marina González Marrero*
Juan de Dios Ortúzar**

DOCUMENTO DE TRABAJO 2004-12

* Universidad de La Laguna. Departamento de Análisis Económico.
** Pontificia Universidad Católica de Chile

PREFERENCE HETEROGENEITY AND WILLINGNESS TO PAY FOR TRAVEL TIME SAVINGS

Francisco Javier Amador
Rosa Marina González

Instituto Universitario de Desarrollo Regional
and
Departamento de Análisis Económico
Universidad de La Laguna, Tenerife, Spain
e-mail: famador@ull.es, rmglzmar@ull.es.

Juan de Dios Ortúzar

Departament of Transport Engineering
Pontificia Universidad Católica de Chile, Santiago, Chile
e-mail: jos@ing.puc.cl

ABSTRACT

We examined different model specifications to detect the presence of preference heterogeneity in a mode choice context. The specification that worked best allows for both systematic and random variations in tastes. Using parameters obtained at the individual level through Bayesian inference methods, subjective values of travel time (SVT) and expected individual compensated variation were derived and aggregated to obtain measures of social welfare. Results suggest that the benefit measures, both at the individual and at the social level, are sensitive to preference heterogeneity assumptions. SVT and welfare changes derived from travel time reductions could be underestimated if the traditional assumption of taste homogeneity is made (we detected differences up to 30% in both types of measures). We also obtained an empirical value for the error made when evaluating changes in social welfare using an approximation of the expected individual compensated variation (expressed as a function of individual SVT) rather than its exact expression.

Key Words: Preference heterogeneity, subjective value of travel time, compensated variation, random parameters logit, Bayesian methods.

JEL: **D61, C25, R41, C11**

1. INTRODUCTION

Measures of the impact of different transport polices on social welfare have traditionally been obtained from fairly simple models. If tastes are assumed to be homogeneous it is possible to derive a single willingness-to-pay (WTP) value for a fictitious average individual. This assumption can be too restrictive as WTP may vary from one person to another depending not only on observable social and economic characteristics, but also on unobserved variables or attributes which are difficult to measure. For this reason it is important to study the distribution of preferences in the population to obtain more accurate measurements.

Advances in simulated estimation techniques have enabled analysts to use increasingly complex models that allow one to define broader behavioural patterns (Train, 2003). However, these models are still rarely applied to evaluation studies and a consensus on the correct way to interpret their results has not yet been reached (Hensher and Greene, 2003; Sillano and Ortúzar, 2004).

This paper has two objectives. First, to analyse individual preference heterogeneity using different approximations, and second, to obtain (both at the individual and social level) and compare, the benefit measures derived from the various approximations used. To address the first objective we considered a battery of models consistent with different hypotheses on individual behaviour. This enabled us to determine whether preference heterogeneity existed. To capture the heterogeneity of individual preferences we used two approaches. The first is a specification in which each attribute parameter is a function of observed socio-economic characteristics of the individuals (i.e. age, sex, income, vehicle ownership). This allows one to model *systematic taste variations* and to identify sources of variability for different WTP measures (Rizzi and Ortúzar, 2003). The second approach attempts to capture *random taste variations* through the specification of a Mixed Logit model. This allows one to obtain both population and individual parameters, the latter by combining simulated maximum likelihood estimates of the former with Bayesian inference methods (Revelt and Train, 1999). Both approaches can also be used in a single model allowing us to incorporate non-observed heterogeneity as well as systematic variations in preferences.

To address the second objective, individual WTP measures (i.e. the subjective value of travel time savings, SVT, and the expected compensated variation, CV) were calculated from a specification allowing for taste variations and from a standard one imposing preference homogeneity, for several hypothetical scenarios. These individual welfare measures were aggregated over the population following the approach of Gálvez and Jara-Díaz (1998), in order to obtain both an *approximate* and an *exact* monetary measure of social welfare. As far as we know, this is the first time that measures of social welfare are derived using individual level parameters. Although the approach is valid even when there are income effects, we applied it in a constant marginal utility of income context since we did not detect income effects in our application. Finally, the results from the different specifications were compared.

The rest of the paper is organised as follows. Section 2 presents the theoretical foundations of discrete choice models and the econometric formulations that can be posed to identify different types of heterogeneity. The next section describes the data bank used for estimating the models. Section 4 presents and discuss the estimated model results. Section 5 makes a presentation of the theoretical framework underpinning the different measures of welfare and compares the resulting values for each of the estimated models. Finally, section 6 summarises our main conclusions.

2. THEORETICAL BASIS AND ECONOMETRIC FORMULATIONS

Most discrete choice models are based on Random Utility Theory (Domencich and McFadden, 1975; Ortúzar and Willumsen, 2001), which postulates that an individual *q* associates a utility (U_{iq}) to each alternative *i* and chooses the alternative with maximum utility. As the analyst is not aware of all the attributes and individual tastes that govern behaviour and as there are also measurement errors involved, he needs to view utility as a stochastic variable made up of the sum of two components:

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

where V_{iq} is the indirect utility function *conditional* on alternative *i* and depends on the attributes which can be measured by the analyst; ε_{iq} is a stochastic component that reflects everything that the modeller can not measure or observe, and helps to explain what would be otherwise considered irrational behaviour by the individual.

The expression used for the deterministic component of utility is frequently a linear function in both the attributes and parameters, that is:

$$V_{iq} = \sum_{k=1}^{K} \beta_{ik} x_{ikq} \tag{2}$$

where x_{ikq} is the value of the k-th attribute of alternative i for individual q and β_{ik} the parameter associated to this attribute; these parameters are usually considered constant for all individuals although they can vary across alternatives.

A microeconomic foundation underlying a linear formulation of V_{iq} when modelling mode choice can be found, for example, in Bates (1987). In this case, after eliminating the elements that do not vary when choosing mode we get for a given individual that:

$$V_{iq} = \alpha_i - \lambda C_{iq} - \psi_i T_{iq} \tag{3}$$

where C_{iq} and T_{iq} are the cost and time of travelling by mode i for individual q, ψ_i is the marginal utility of reducing the minimum travel time by mode i and λ is the marginal utility of income, given as usual by: $\lambda = -\partial V_{iq} / \partial C_{iq}$.

A conventional form of reflecting heterogeneity of preferences has been to introduce interactions between alternative attributes and individual socio-economic characteristics such as sex, age, income level, or even specific trip characteristics (e.g. purpose, travel conditions), see Pollak and Wales (1992), Revelt and Train (1998), Ortúzar and Willumsen (2001). Here the parameter of each attribute (β_{ik}) is allowed to be a function of the individual's observed socio-economic characteristics and, as mentioned above, this allows us to detect systematic variations in tastes.

However, on many occasions individual information is not available, or tastes may vary with characteristics that are difficult to measure or cannot be observed. In these cases (2) can be generalised to consider heterogeneity specifying random parameters for each individual. Thus, the utility of alternative i for individual q would be:

$$V_{iq} = \beta_q x_{iq} = (b + \eta_q) x_{iq} \tag{4}$$

where β_q is now a vector of coefficients for each individual q that varies randomly with tastes and can be expressed as the sum of a population mean (b) and individual deviations from the average value for the population (η_q).

One disadvantage of specifying random parameters is that information is not provided about factors determining taste variations across individuals. So, it could be convenient to use a specification that contains both interactions and random parameters.

If a fixed parameters model allowing for systematic variation of tastes is considered ($\beta_q = \beta, \forall q = 1,...,N$) and the error terms ε distribute iid Gumbel, we obtain the well-known Multinomial Logit (MNL) model (Domencich and McFadden, 1975). The probability of individual q choosing alternative i is given in this case by:

$$L_{iq} = \frac{\exp(\varphi \beta X_{iq})}{\sum_{j=1}^{J} \exp(\varphi \beta X_{jq})} \tag{5}$$

where φ is a scale factor inversely proportional to the variance of the error term (i.e. $\varphi = \frac{\pi}{\sqrt{6}\sigma}$). In standard model applications the scale factor is normalized (i.e. taken as one) as it cannot be estimated separately from the vector of taste parameters β.

On the other hand, if we allow for random variations in the taste parameters (β_q) and use the same distribution for the error terms (ε_{iq}), we get a Mixed Logit (ML) random parameters model, in which the utility of alternative i is given by the following expression (Train, 2003):

$$U_{iq} = \beta_q x_{iq} + \varepsilon_{iq} \tag{6}$$

where ε_{iq} ~ Gumbel (0, σ^2) and $\beta_q \sim f(\beta_q | b, \Sigma)$; f is a general density function and b, Σ represent the mean and covariance characterising its distribution in the population. The model can also be generalised to consider that each individual faces a sequence of T choices $y_q = (y_{1q},..., y_{Tq})$, allowing us to model the correlation between different choice situations faced by an individual correctly, as in the case of stated preference data or panel data (Train, 2003).

Since β_q is unknown, the probability of the individual's choice conditional on the distribution of the population parameters (i.e. the mixed logit choice probability) is :

$$P(y_q | b, \Sigma) = \int L(y_q | \beta_q) f(\beta_q | b, \Sigma)\, d\beta_q \tag{7}$$

where $L(y_q | \beta_q)$ is the probability of individual q's observed choice conditional on β_q, and coincides with (5). McFadden and Train (2000) show that any discrete choice model derived from random utility maximization can be approximated arbitrarily closely by a ML model with the appropriate specification of $f(\beta_q | b, \Sigma)$.

As (7) does not have a closed form the probability is approximated numerically through simulation. In particular, draws of β_q are taken from $f(\beta_q | b, \Sigma)$. For each draw, $L(y_q | \beta_q)$ is calculated and the results are averaged over draws. Once the simulated probabilities are obtained they are used to construct a simulated log-likelihood (SLL) function which is maximised. The maximum SLL estimates of b and Σ define the frequency distribution of the individual parameters β_q for the population (Train, 2003).

Different drawing techniques can be used to reduce the simulation variance and to improve the efficiency of the estimation. In this study parameters were estimated from 125 Halton draws using the GAUSS code developed by Train, Revelt and Ruud[1]. This method provides better accuracy with fewer draws than the typical pseudo random draws and requires less computational time (Bhat, 2001)[2].

Once b and Σ have been estimated, the *classical approach*, described by Revelt and Train (1999) can be followed[3] to obtain point estimates for each β_q. This involves the conditioning of (b, Σ) to the individual choices and this is done as follows. Let $h(\beta_q | y_q, b, \Sigma)$ denote the density of β_q conditional on the individual's sequence of choices y_q and the population parameters b, Σ. By Bayes' rule we have:

$$h(\beta_q | y_q, b, \Sigma) = \frac{P_q(y_q | \beta_q) f(\beta_q | b, \Sigma)}{P_q(y_q | b, \Sigma)} \tag{8}$$

---

[1] The code may be downloaded from Prof. Kenneth Train's web page: http:\\elsa.Berkeley.EDU\~train\
[2] He shows that in higher dimension integrals, 125 Halton draws provide the same level of accuracy as 2000 pseudo random draws and require much less time for convergence. However the field continuous to evolve and now it appears that Sobol sequences are even better, in particular for problems with many dimensions (Silva and Garrido, 2003).
[3] Instead of using this *classical* method to estimate individual-level parameters, the *Hierarchical Bayes* method (Albert and Chib, 1993; McCulloch and Rossi, 1994; Allenby and Rossi, 1999) can be used. Huber and Train (2001) and Godoy and Ortúzar (2004) investigate the similarity of *classical* and Bayesian results for the mixed logit model with not entirely consistent findings. Train (2001) analyses the convenience of the two methods in terms of programming and computing time, depending on the specification of the model. An application of both methods can be found in Sillano and Ortúzar (2004).

The conditional expectation β results from integrating over the domain of β_q. This integral can be approximated by simulation, averaging weighted draws β_q^r from the population density function $f(\beta_q \mid b, \Sigma)$. The simulated expectation *SE* is given by:

$$SE(\beta_q \mid y_q, b, \Sigma) = \frac{\sum_{r=1}^{R} \beta_q^r P_q(y_q \mid \beta_q^r)}{\sum_{r=1}^{R} P_q(y_q \mid \beta_q^r)} \tag{9}$$

where R is the number of draws. Revelt and Train (1999) also propose, but do not apply, an alternative simulation method to condition individual level choices. Consider the expression for $h(\beta_q \mid y_q, b, \Sigma)$ in (8). The denominator is a constant value since it does not involve β_q, so a proportionality relation can be established as:

$$h(\beta_q \mid y_q, b, \Sigma) \propto P_q(y_q \mid \beta_q)\ f(\beta_q \mid b, \Sigma) \tag{10}$$

From this relation, and using the maximum SLL estimates of the population parameters (b, Σ) obtained at the first stage, we can draw observations of the *posterior* distribution $h(\beta_q \mid y_q, b, \Sigma)$ using the Metropolis-Hastings algorithm (Chib and Greenberg, 1995). This process is used to take repeated draws from the *prior* distribution $f(\beta_q \mid b, \Sigma)$ which will be taken as valid values of β_q if they contribute to increase the SLL. Successive draws of β_q generally provide an increasingly better fit of the model to the data, until such time as improvements are no longer possible. When this occurs we consider that the iterative process has converged. After a number of *burn-out* iterations to ensure that a steady state has been reached (typically a few thousand), one every ten sampled values generated in the process are stored (to avoid serial correlation). From these, a sampling distribution for $h(\beta_q \mid y_q, b, \Sigma)$ can be built and inferences about the mean and covariance values can be obtained (Sawtooth Software, 2000). This procedure was used in this paper[4].

[4] For estimating individual parameters the procedure was coded in WinBUGS, performing 100 000 burn out iterations prior to the 20 000 used to build the posterior distribution. WinBUGS was developed by the MRC Biostatistics Unit at the University of Cambridge and the Imperial College of Medicine at St Mary's, London. It can be downloaded from the site: http:\\www.mrc-bsu.cam.ac.uk\bugs\welcome.shtml

3. DATA

The information used in this paper came from a survey carried out to students of the Faculty of Economic and Business Sciences, University of La Laguna (Spain), in the week 22-26 May 2000. Revealed preference data for the choice of transport mode to the university was obtained, and also under what conditions the choice was made.

The design of the original questionnaire was evaluated through focus groups using randomly selected students who were later to form part of the survey sample. This allowed us to detect possible ambiguities and ensure that, as far as possible, the questions would be understood by the potential respondents. The information obtained from the survey was screened to exclude students who were captive to a given mode. The final sample consisted of 494 undergraduate and graduate students from Economics and Business Administration. Of these, 204 were male and 290 female.

Table 1 shows the frequency of choice and the availability of each option, by sex. We can observe that most students drive to the Faculty (50% of the women and 58% of the men). When other modes are analysed one can detect a higher difference in profile between men and women. For men, the second most preferred modes are bus and on foot (around 13% in both cases), followed by travelling as a passenger in a private car (12%). For women the runner up option is to travel as car passengers (22%) followed by the conventional bus (10%). The data also shows that practically all the travellers with access to a private vehicle use it (92% of women and 88% of men).

Table 1: Choice and Availability by Sex

Transport Mode	Choice (%)		Availability (%)	
	Men	Women	Men	Women
Car-driver	58.33	50.00	66.18	54.14
Car-passenger	11.76	22.41	38.24	48.97
Bus	12.75	10.00	82.35	83.10
University shuttle bus	2.94	8.97	13.24	20.00
Motorcycle	0.98	0.34	2.94	1.38
On foot	13.24	8.28	31.86	30.34

Finally, a comparison of the social and economic variables included in the questionnaire (family income level, student allowance, possession of private vehicle, etc.) shows that the profiles are very similar for men and women (see Table 2).

Table 2: Monthly Family Income and Student Allowance by Sex

Variable	Men (%)	Women (%)
Family income (pts[a]./month)		
Less than 75,000	15.69	20.34
Between 75,000 and 150,000	28.92	25.86
Between 150,000 and 250,000	25.49	26.21
Between 250,000 and 400,000	12.75	7.59
More than 400,000	3.92	6.55
No answer	13.24	13.45
Student allowance (pts./month)		
No allowance	39.71	45.52
Between 2,000 and 10,000	9.80	6.90
Between 10,000 and 22,000	10.78	10.00
Between 22,000 and 32,000	5.88	6.21
Between 32,000 and 45,000	9.80	12.41
Between 45,000 and 70,000	7.35	2.76
Over 70,000	7.84	4.48
No answer	8.82	11.72

[a] At the time of the study 1 Euro was equal to 166 pts.

4. MODEL ESTIMATION AND RESULTS

All the reported models were specified with linear utility functions following (2) and (6). The explanatory variables considered were *Travel time* and *Cost* and, in the case of Bus, *Waiting time*, defined as the average time between departures from the bus station[5]. Table 3 shows the mean and standard deviation of the explanatory variables. All variables were specified with generic parameters; a specific modal constant was also included for each alternative except for Car-driver which was taken as reference.

We first estimated a Multinomial Logit model (MNL-1) imposing homogeneity on population tastes (see Table 4). The results can be considered generally acceptable, not only because the signs of the estimated coefficients are intuitively correct, but also because most *t* statistics are satisfactory. As *Waiting time* had a correct sign, it was kept in the model in spite of its rather low significance (Ortúzar and Willumsen, 2001).

Different specifications were then tested to see if there was evidence of systematic heterogeneity in preferences, based on the variables sex, age, possession of a vehicle, family income level, student’s allowance, distance and work status.

[5] Travel cost is expressed in pesetas, travel time and waiting time in minutes.

Table 3: Descriptive Statistics of the Explanatory Variables

Variable	Mean	Standard Deviation
Time Car-driver	24.17	24.41
Cost Car-driver	205.38	145.84
Time Car-passenger	27.46	21.35
Cost Car-passenger	92.29	124.52
Time Bus	50.98	27.69
Cost Bus	195.07	160.79
Waiting time Bus	23.09	25.63
Time University bus	17.95	9.82
Cost University bus	26.53	10.03
Time Motorcycle	12.20	10.92
Cost Motorcycle	98.5	28.48
Time On foot	26.22	17.77

Dummy variables defined for each of these variables were made to interact with *Travel time (TT)*, *Waiting time (WT)* and *Cost (C)*. We found that individual tastes tend to vary little with observed socio-economic features across the sample. The only significant interaction was that between *Travel time* and the dummy Sex_q which takes the value of one for men and zero otherwise (model MNL-2). The deterministic utility in this case was:

$$V_{iq} = \beta_i + (\beta_{TT} + \beta_{Sex} Sex_q) TT_{iq} + \beta_C C_{iq} + \beta_{WT} WT_{iq} \qquad (11)$$

where β_i is the specific constant of the alternative. Also recall that (- β_{Cost}) is the marginal utility of income (λ). Equation (11) implies that the hypothesis that men and women have different perceptions of time is accepted, so model MNL-1 is rejected in favour of MNL-2 (this was confirmed by a likelihood ratio test). This finding is consistent with the descriptive analysis of the data which revealed clearly different choice profiles for men and women.

We also specified a model with a cost-squared variable to check the existence of income effects (Jara-Díaz and Videla, 1989), but the variable was not significant. This allows us to state confidently that λ may be considered independent of the income level.

Table 4: Maximum Likelihood Estimation of MNL and ML Models

Coefficients (t-test)		*MNL-1*	*MNL-2*	*ML-1*	*ML-2*
Travel time	*Mean*	-0.0460	-0.0584	-0.0706	-0.0792
		(-4.50)	(-4.60)	(-2.89)	(-3.05)
	Spread[1] *(s)*	-	-	0.0800	0.0742
		-	-	(2.07)	(1.87)
*Sex*Travel time*		-	0.0258	-	0.0248
		-	(1.90)	-	(1.66)
Cost		-0.0031	-0.0031	-0.0033	-0.0032
		(-2.30)	(-2.30)	(-2.65)	(-2.58)
Waiting time		-0.0102	-0.0092	-0.0120	-0.0109
		(-1.30)	(-1.20)	(-0.96)	(-0.85)
Specific Constants	*Car-Passenger*	-2.6780	-2.7220	-2.8024	-2.8121
		(-8.50)	(-8.60)	(-8.37)	(-8.38)
	Bus	-2.0300	-2.0480	-1.9718	-1.9992
		(-6.00)	(-6.10)	(-4.71)	(-4.77)
	University Bus	-1.9010	-1.9220	-1.8949	-1.9058
		(-5.20)	(-5.30)	(-4.63)	(-4.66)
	Motorcycle	-1.7400	-1.7480	-1.8135	-1.7816
		(-2.00)	(-2.00)	(-1.70)	(-1.59)
	On Foot	-1.7110	-1.7390	-1.5639	-1.5934
		(-4.80)	(-4.90)	(-3.31)	(-3.37)
Sample Size		494			
Log Likelihood		-224.8168	-223.0512	-223.4809	-222.2752
$\bar{\rho}^2$		0.077	0.084	0.083	0.088

[1] S*pread* (s) is the distance between the mean (m) and the extreme of the uniform distribution; thus the interval of the distribution is defined as [m-*s*, m+*s*].

A range of hypotheses was considered to test for random taste variations. One assumed that only the perception of time varied randomly across travellers, a second considered that the marginal utility of income varied and, a third combined the previous two. Moreover, different distributions were considered for each random parameter: normal, uniform and triangular.

We found that the *Travel time* parameter was the only one presenting a significant random variation over the population (ML-1); the *t*-test for the spread (*s)* of this parameter indicates that it is statistically significant at the 95% confidence level. This reinforces the hypothesis that tastes vary (which had been detected to some extent by model MNL-2), but it does not allow us to determine whether or not there is a relation between tastes and the observed characteristics of the individuals. Finally, as model MNL-2 suggested the existence of an interaction between sex and time we decided to check whether there were additional sources of heterogeneity (random, or due to individual idiosyncrasies). The results of model ML-2 indicate that this is indeed the case as the spread *s* was again significantly different from zero with a confidence interval of over 90%. Furthermore unlike ML-1, model ML-2 has the advantage of not allowing the *Travel time* parameter to take positive values. This means that in this case no individual has a parameter with an incorrect sign, which is an excellent result (see the discussion in Sillano and Ortúzar, 2004). Therefore, although there are no significant gains in terms of fit our proposed specification to capture preference heterogeneity is one introducing a random uniform parameter for *Travel time* and simultaneously allowing for the interaction of time and sex (ML-2). We considered this model to derive individual parameters.

Based on the population parameters for model ML-2, the Revelt and Train (1999) secondary approach was applied to produce individual point estimates for the random parameters $\beta_{Time,q}$. This set of values together with the rest (non-random) of the parameters is referred to as model ML-3. The log-likelihood function evaluated for this model yields the value –194.5, which is substantially better than the values obtained for the models in Table 4. This confirms that a model with individual parameters is able to achieve a significantly better fit to the data as the individual parameters allow to characterise the likelihood function more accurately (Godoy and Ortúzar, 2004).

5. WELFARE MEASURES

5.1 Individual Willingness to Pay

A typical application of random utility models is to estimate the subjective value of travel time savings (SVT). The SVT is defined as the maximum amount an individual is willing to pay to reduce travel time by one unit[6] and, given a linear indirect utility formulation this is equal to the ratio of the time and cost coefficients (Gaudry *et al*, 1989). Note that this is, by definition, a marginal WTP measure.

Table 5 presents SVT values obtained for each specification presented in this paper. The results indicate, first, that the SVT values derived from a model with homogeneous preferences (MNL-1) can be similar to those obtained when systematic variations in tastes are considered (MNL-2); however if travel time tastes are allowed to vary randomly, significant differences appear (i.e. up to 40% increase in SVT). This suggests that using a restrictive specification may lead to an underestimation of the value of travel time savings.

Table 5: Subjective Values of Travel Time[a] (pts[b]/min)

	Men	*Women*	*Mean*
MNL-1	-	-	14.9 (14.3 – 15.6)
MNL-2	10.4 (10.0 – 10.8)	18.7 (17.9 – 19.4)	15.3[c]
ML-1	-	-	21.4 (20.4 – 22.4)
ML-2	17.0 (16.4 – 17.6)	24.7 (23.7 – 25.9)	21.5[c]

[a] Confidence intervals for the SVT at the 95% level are presented in parenthesis following Armstrong *et al*. (2001).
[b] Pesetas of the year 2000.
[c] This figure is a weighted average considering that the sample is composed of 204 men and 290 women.

It is interesting to note that international experience suggests that this is not a general conclusion but depends on the nature of the data and specifications used in each study. For example, Hensher (2001a, 2001b) also concludes that more restrictive models tend to underestimate the value of time; however, other authors have found no significant differences between the values produced by different models (Train, 1998; Carlsson, 2003), and in some cases even lower SVT values have been obtained when Mixed Logit (Algers *et al*., 1998) or more flexible models than the MNL are specified (Gaudry *et al*, 1989). Finally, Alpizar and Carlsson (2001) found that the SVT could be underestimated or overestimated depending on the chosen mode.

[6] For a theoretical review of the time allocation models necessary to derive subjective values of time, see González (1997); an early discussion about the influence of model structure and specification on *SVT* can be found in Gaudry *et al* (1989).

One possible explanation for these empirically observed discrepancies is the re-scaling that all parameters undergo when we move from a fixed specification to one where some parameters are allowed to vary randomly (Sillano and Ortúzar, 2004). The random parameter specification explains part (η_q) of the non-observed component of utility (V_{iq}) – see equation (4). This way, the variance (σ^2) of the iid Gumbel (ε_{iq}) error component in the ML model is smaller than in the MNL model producing an increase in the associated scale factor (φ). Re-scaling occurs as long as the scale factor directly affects the estimation of all parameters.

In practise, if all parameters were re-scaled in the same proportion the SVT should not be affected by changing the specification. The empirical evidence shows, however, that not all parameters are re-scaled by the same magnitude. This could be due to a problem similar to what happens when a relevant variable is omitted, as the parameters reflecting variations in population tastes are included in the ML but are omitted in the MNL. Thus, depending on the variables included in the model, the functional form chosen for the indirect utility function and the nature of the data, a fixed parameters model may lead to over/under estimates of the true values of time.

Now, using the point estimates of $\beta_{Time,\,q}$ (ML-3) it is possible to provide additional valuable information on the distribution of the SVT over the population. For instance, in our sample, the median SVT (24 pts/min) is slightly higher than the mean (which is equal to the mean for ML-2), and the range is between 1.2 and 30 ptas/min. More detailed analyses on individual results could be made (e.g. cluster analysis), but these are beyond the scope of this paper.

When there are major savings in travel time, perhaps inducing significant substitution among alternatives, simply multiplying the SVT by the variation in travel time is no longer appropriate for evaluating individual welfare changes[7]. To this extent, a Hicksian measure of individual welfare like the compensating variation (CV_q) could be computed to obtain an exact monetary measure. For a reduction in travel time, CV_q measures the maximum individual WTP for the saving and can be defined as the value CV satisfying[8]

$$U(M, c, t^0, \varepsilon) = U(M - CV, c, t^1, \varepsilon) \tag{12}$$

[7] This implies assuming that individuals will choose the same alternative before and after the change.
[8] Alternatively, the equivalent variation (EV_q) can be used. For a reduction in travel time, EV_q measures the minimum amount the individual is willing to accept to forgo the time saving.

where *0* and *1* denote the situation before and after the change and $U(.)$ is the individual's *unconditional* indirect utility function:

$$U(M,c,t,\varepsilon) = \underset{i}{Max}\ V_i + \varepsilon_i \tag{13}$$

The problem of calculating the compensating variation $CV = CV(M,c,t^0,t^1,\varepsilon)$ in a random utility framework is that it is a random variable and, in general, there is no closed solution[9] for it. However, when the marginal utility of income is constant and the random component of utility is Generalized Extreme Value (GEV) distributed, an explicit *logsum* form for the *expected* compensating variation $E[CV_q]$ can be obtained (McFadden, 1978). In the particular case of ε_i distributed iid Gumbel, we get:

$$E\left[CV_q\right] = \frac{1}{\lambda_q}\left[\ln \sum_{i \in A(q)} \exp(V_{iq}^0) - \ln \sum_{i \in A(q)} \exp(V_{iq}^1)\right] \tag{14}$$

where V_{jq}^0 and V_{jq}^1 are the indirect utility functions conditional on mode *j* for individual *q* in the initial and final situations respectively (Williams, 1977; Small and Rosen, 1981). It should be noted that, since the marginal utility of income is constant, there are no income effects so (a) Hicksian welfare measures, equivalent and compensating variation, coincide with the traditional Marshallian consumer surplus variation, (b) their expected values are given by (14) (McFadden, 1981; Hanemann, 2001) and (c) expression (14) provides an *exact* measure of individual welfare.

On the other hand, when there are no income effects but one has a ML specification with at least one random parameter, the calculation of $E[CV_q]$ requires integration $\int E\left[CV_q(\beta)\right] f(\beta \mid b,\Sigma)\, d\beta$, given a known distribution of β over the population (Train, 1998), or simulation (e.g. Breffle and Morey, 2000). However, since the Revelt and Train (1999) approach provides point estimates of the individual level parameters β_q, an alternative way to evaluate $E[CV_q]$ is to use these point estimates directly [10]. This is the procedure that we followed in this paper.

---

[9] The calculation of $E[CV_q]$ is in most cases analytically intractable. Various approximations and simulation methods have been proposed to calculate it when there are income effects (e.g. McFadden, 1999; Herriges and Kling, 1999), but it has been only recently that an exact solution (in the form of a one finite dimension integral) has been derived for GEV random utility models (Karlström, 2001).

Finally, it is worth asking whether there is a relationship between $E[CV_q]$ and SVT_q. Following Jara-Díaz (1990), Gálvez and Jara-Díaz (1998) derived an approximation for $E[CV_q]$ that is a function of SVT under some linearity assumptions[11]. The resulting expression when a project only involves time savings is given by

$$E\left[CV_q\right] \approx SVT_q \sum_i \overline{P}_{iq} (t_i^0 - t_i^1) = SVT_q \, TTS_q \tag{15}$$

where $\overline{P}_{iq} = \frac{P_{iq}^0 + P_{iq}^1}{2}$ is the average probability of individual q choosing alternative i before and after the change. Then, $TTS_q = \sum_i \overline{P}_{iq} (t_i^0 - t_i^1)$ is an approximation of the expected travel time saved by individual q.

5.2 Monetary Measurement of Social Welfare Changes

Once the individual WTP values have been estimated we face the problem of aggregating them into a single value that can act as reference in the decision making process. Following the social welfare approach, Gálvez and Jara-Díaz (1998) provide a general framework to deal with the social appraisal of projects financed with public funds. Let W be a social welfare function that depends on the utility of every individual or group q,

$$W = W(U_1, ..., U_q, ..., U_n) \tag{16}$$

Individual utility U_q is a direct function of goods consumption X_q; the latter depends on goods prices P, goods characteristics Q and individual income I_q. Thus:

$$U_q = U_q\left\lfloor X_q(P, Q, I_q)\right\rfloor = V_q(P, Q, I_q) \tag{17}$$

where V_q is an indirect utility function.

Assuming that a monetary measure of welfare change dB_q for each individual (or group) has been obtained, Gálvez and Jara-Díaz (1998) show that the variation in social welfare dW after a project is given by:

---

[10] Von Haefen (2003) has recently developed an alternative procedure to evaluate (14) using conditional information on individual tastes.

[11] In particular, choosing a linear trajectory to solve the integral that yields the CV they assume that the probability of being chosen varies linearly with travel time and that the indirect utility function adopts a linear form.

$$dW = \sum_q \frac{\partial W}{\partial U_q} \frac{\partial V_q}{\partial I_q} dB_q = \sum_q \Omega_q \lambda_q dB_q \tag{18}$$

where Ω_q reflects the importance that society assigns to the welfare of each individual (i.e. a "social weight") and λ_q is, as before, the individual marginal utility of income. Therefore, the social welfare variation can be expressed as a weighted sum of the monetary measures of benefit by all individuals.

As dW is expressed in social utility units, a "social conversion" factor λ_s is needed to convert it into money terms dB, such that:

$$dB = \frac{dW}{\lambda_s} = \frac{1}{\lambda_s} \sum_q \Omega_q \lambda_q dB_q \tag{19}$$

The approach requires defining a set of social weights and determining the value of λ_s. In what follows we assume that $\Omega_q = 1$, a neutral scheme assuming that all individuals have the same social weight. On the other hand, Gálvez and Jara-Díaz (1998) propose using the ratio between the social loss due to tax paying and the tax bill to calculate λ_s. For equal social weights this results in a weighted average of the individual marginal utilities of income, using tax proportions as weights.

Thus, (19) provides an analytical framework for the social appraisal of projects in the general case that all preference parameters vary in the population, This is valid even when the cost parameter, and thus the marginal utility of income, varies randomly over the population. However, since in our application the marginal utility of income is constant it is convenient to analyse what happens to (19) in this case. Under these circumstances, it can be shown that $\lambda_s = \lambda$ if the full costs of the project are borne by its beneficiaries. In such a case the social benefit is equal to the direct summation of the monetary measures of individual welfare[12]:

$$dB = \sum_q dB_q \tag{20}$$

[12] Gálvez and Jara-Díaz (1998) point out that when the marginal utility of income varies between income strata, calculating a monetary measure of social welfare change by simply adding dB_q over the population is a clearly regressive aggregation criterion, as it involves assigning greater social weights to individuals with a higher income level.

Thus, to evaluate *dB* the only remaining action is to choose a measure for dB_q. Gálvez and Jara-Díaz (1998) propose to use as an *approximate* measure of dB_q the approximate $E[CV_q]$ given by expression (15), such that:

$$dB \approx \sum_q SVT_q TTS_q = \sum_q \frac{|\beta_{Time,q}|}{\lambda_q} TTS_q \tag{21}$$

A particular case of this expression is obtained when the marginal utility of income is constant and time preferences are homogeneous in the population so that $\beta_{Time,q} = \beta_{Time}$. In this case a single value of time can be used to obtain measures of social welfare. But, if there are variations in travel time tastes across the population, using a single time value would lead to an incorrect measurement of social welfare.

Alternatively, *dB* could be derived from the *exact* expression of $E[CV_q]$ given by (14). Thus, when the marginal utility of income is constant one has that:

$$dB = \frac{1}{\lambda} \sum_q \left[\ln \sum_{j \in A(q)} \exp(V_{jq}^0) - \ln \sum_{j \in A(q)} \exp(V_{jq}^1) \right] \tag{22}$$

As a result, two empirical issues related to the calculation of *dB* arise. One is to evaluate the sensitivity of *dB* to the assumptions on individual preferences for travel time. The second is to quantify the magnitude of the error produced when evaluating changes in social welfare using the dB_q approximation proposed by Gálvez and Jara-Díaz (1998) instead of its exact value.

To address the first issue we compare the values of *dB* derived using the *exact logsum* measure of $E[CV_q]$ resulting from models MNL-1, MNL-2 and ML-2 (see section 4). For this, six hypothetical scenarios were considered in which travel time by car-driver (in what follows car) and by bus, were reduced by 10%, 30% and 50% respectively. The corresponding values of *dB* are presented in Table 6.

In Table 7 we present the percentages by which the more restrictive models tend to underestimate the monetary measures of change in social welfare. As it occurred when we analysed the SVT, our results lead us to conclude that the social welfare measure is indeed sensitive to assumptions concerning the behaviour of individual preferences.

Table 6: Social Welfare Changes Using Exact Logsum Expression for dB_q

	Travel time saving by car		
	10%	30%	50%
MNL-1	9366.5	28371.2	47652.6
MNL-2	9746.3	29548.4	49657.3
ML-3	13901.7	42176.4	70839.5
	Travel time saving by bus		
	10%	30%	50%
MNL-1	3450.9	12501.2	25310.4
MNL-2	3475.0	12899.5	26922.5
ML-3	4016.5	16022.7	36331.2

Table 7: Differences in Social Welfare Changes with Respect to Model ML-3 (%)

	Travel time saving by car		
	10%	30%	50%
MNL-1	-32.62	-32.73	-32.73
MNL-2	-29.89	-29.94	-29.90
	Travel time saving by bus		
	10%	30%	50%
MNL-1	-14.08	-21.98	-30.33
MNL-2	-13.48	-19.49	-25.90

Analogously, the lowest values for *dB* are obtained when homogeneity in preferences is imposed (MNL-1), yielding values up to 32% smaller than those resulting from ML-3. Moreover, the percentage difference tends to increase with the reductions in travel time, although the differences remain almost constant for the reductions in travel time by car. Note that if only systematic taste variations are considered (MNL-2), the measures of social welfare are still underestimated although the differences are slightly smaller.

Finally, to evaluate the differences in social welfare change obtained with the approximation of Gálvez and Jara-Díaz (1998) with respect to the exact expression for $E[CV_q]$, the percentage differences between both measures were calculated. The results are presented in Table 8.

Table 8: *dB* approximation error (%)

	Travel time saving by car		
	10%	30%	50%
MNL-1	-0.02	-0.12	-0.25
MNL-2	-0.02	-0.16	-0.28
ML-3	0.07	-0.14	-0.39
	Travel time saving by bus		
	10%	30%	50%
MNL-1	0.28	2.35	5.41
MNL-2	0.37	3.23	7.31
ML-3	0.88	6.09	12.48

As one would expect, the quality of the approximation becomes worse as the travel time savings increase yielding errors up to 12% in the case of bus. Notwithstanding, the errors in the case of car are practically negligible. This result is due to the fact that, unlike the bus, the probability of choosing car does not vary significantly between the different scenarios[13].

On the other hand, when analysing the savings in travel time by bus we found that the approximate measure overestimated the increase in social welfare, although significant differences between both measures were only apparent when large savings in travel time accrued. In general, the magnitude of the errors are relatively small, meaning that the Gálvez and Jara-Díaz (1998) approximation may be considered valid.

[13] The probabilities of choosing each alternative (car and bus) were calculated for the different scenarios. We found that the percentage of probability change varied between 0.75 and 3.30 % for the car and between 13 and 132 % for the bus.

6. CONCLUSIONS

We compared different discrete choice model specifications to detect the presence of taste variations in our sample. We found significant deterministic differences between the preferences of men and women in relation to travel time. However, our results also pointed out to the existence of other sources of heterogeneity in preferences which are of a random nature.

We also derived social benefit measures in a context of heterogeneous preferences. First, to allow for tastes variation, we followed an estimation procedure to obtain individual ML parameter estimates proposed by Revelt and Train (1999), and derived welfare measures (SVT and expected compensating variation) at the individual level. Since the representative individual approach is no more valid when preference are heterogeneous, an aggregation criteria must be followed in order to obtain a measure of social welfare. We suggested that in this context the social welfare approach developed by Gálvez and Jara-Díaz (1998) could provide the necessary theoretical framework to calculate measures of social welfare and we applied it, being the first time that measures of social welfare are calculated from individual level parameters.

The empirical evidence provided in this paper suggests that the benefit measures, both at the individual and social levels, are sensitive to the assumptions about preference heterogeneity. When the traditional assumption of taste homogeneity is made (i.e. a MNL specification) both the subjective value of time and welfare changes derived from travel time reductions could be underestimated. In particular, we detected differences up to 30% in both types of measures.

We also calculated social welfare changes using an approximation for the expected individual compensated variation that can be expressed as a function of individual SVT (Gálvez and Jara-Diaz, 1998). Results derived from simulating several scenarios show that when the approximation is used the magnitude of the errors, although relatively small, depend on: the time reduction simulated, the mode of transport to which it applies and the model specification used. In particular, the approximation seems more valid if the probabilities of choice before and after the policy tested remain approximately equal.

ACKNOWLEDGEMENTS

We wish to thank Mauricio Sillano for his comments and his help in producing the necessary code to use WinBUGS. The third author would like to acknowledge the support of FONDECYT through project 1020981.

REFERENCES

Albert, J. and S. Chib (1993): “Bayesian Análisis of Binary and Polychotomous Response Data”, *Journal of American Statistical Association* 88, 669-679.

Algers, S., P. Bergström, M. Dahlberg, and J.L. Dillen (1998): “Mixed Logit Estimation of the Value of Travel Time”, *Working Paper*, Department of Economics, Upsala University.

Allenby, G.M. and P. Rossi (1999): “Marketing Models for Consumer Heterogeneity”, *Journal of Econometrics* 89, 57-78.

Alpizar, F and F. Carlsson (2001): “Policy Implications and Analysis of the Determinants of Travel Mode Choice: An Application of Choice Experiments to Metropolitan Costa Rica”, *Working Paper in Economics* 56, Department of Economics, Göteborg University.

Armstrong, P., R. Garrido, and J. de D. Ortúzar (2001): “Confidence Intervals to Bound the Value of Time”, *Transportation Research* 37E, 143-161.

Bates, J.J. (1987): “Measuring Travel Time Values with a Discrete Choice Model: A Note”, *The Economic Journal* 97, 493-498.

Bhat, C.R. (2001): "Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model", *Transportation Research* 35B, 677-693.

Breffle, W. and E. Morey (2000): “Investigating Preference Heterogeneity in a Repeated Discrete-Choice Recreation Demand Model of Atlantic Salmon Fishing”, *Marine Resource Economics* 15, 1-20.

Carlsson, F. (2003): “The Demand for Intercity Public Transport: The Case of Business Passengers”, *Applied Economics* 35, 41-50.

Chib, S. and E. Greenberg (1995): “Understanding the Metropolis-Hastings Algorithm”, *American Statistician* 49, 327-335.

Domencich, T. and D. McFadden (1975): *Urban Travel Demand: A Behavioural*

Analysis. North-Holland, Amsterdam.
Gálvez, T. and S.R. Jara-Díaz (1998): "On the Social Valuation of Travel Time Savings", *International Journal of Transport Economics* 25, 205-219.
Godoy, G. and J. de D. Ortúzar (2004): "On the Estimation of Mixed Logit Models", *Statistics and Computing* (under review).
Gaudry, M.J.I., S.R. Jara-Díaz and J. de D. Ortúzar (1989): "Value of Time Sensitivity to Model Specification", *Transportation Research* 23B, 151-158.
González, R.M. (1997): "The Value of Time: A Theoretical Review", *Transport Reviews* 17, 245-266.
Hanemann, W.M. (2001): "Applied Welfare Analysis with Discrete Choice Models". In C. Kling and J. Herriges (eds.), *Valuing the Environment Using Recreation Demand Models*, Edward Elgar, Cheltenham.
Hensher, D.A. (2001a): "The Valuation of Commuter Travel Time Savings for Car Drivers: Evaluating Alternative Model Specifications", *Transportation* 28, 101-118.
Hensher, D.A.: (2001b): "Measurement of the Valuation of Travel Time Savings", *Journal of Transport Economics and Policy* 35, 71-98
Hensher, D.A. and W.H. Greene (2003): "The Mixed Logit Model: The State of Practice", *Transportation* 30, 133-176.
Herriges, J. and C. Kling (1999): "Nonlinear Income Effects in Random Utility Models", *Review of Economics and Statistics* 81, 62-72.
Huber, J. and K. Train (2001): "On the Similarity of Classic and Bayesian Estimates of Individual Mean Partworths", *Marketing Letters* 12, 259-269.
Jara-Díaz, S. (1990): "Consumer's Surplus and the Value of Travel Time Savings", *Transportation Research* 24B, 73-77.
Jara-Díaz, S.R. and J. Videla (1989): "Detection of Income Effect in Mode Choice: Theory and Application", *Transportation Research* 23B, 393-400.
Karlström, A. (2001): "Welfare Evaluations in Non-Linear Random Utility Models with Income Effects". In D. Hensher (ed.), *Travel Behaviour Research: The Leading Edge*, Pergamon Press, Sidney.
McCulloch, R. and P. Rossi (1994): "An Exact Likelihood Analysis of the Multinomial Probit Model", *Journal of Econometrics* 64, 207-240.

McFadden, D. (1978): "Modelling the Choice of Residential Location". In A.

Karlquist, L. Lundquist, F. Snickars and J.W. Weibull (eds.), *Spatial Interaction Theory and Planning Models,* North-Holland, Amsterdam.

McFadden, D. (1981): "Econometric Models of Probabilistic Choice". In C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, The MIT Press, Cambridge, Mass.

McFadden, D. (1999): "Computing Willingness-to-Pay in Random Utility Models". In J. Moore, R. Riezman, and J. Melvin. (eds*.), Trade, Theory and Econometrics: Essays in Honour of John S. Chipman*, Routledge, London.

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

Ortúzar, J. de D. and L.G. Willumsen (2001): *Modelling Transport.* 3rd Edition, John Wiley and Sons, Chichester.

Pollak, R.A. and T.J. Wales (1992): *Demand Specification and Estimation.* Oxford University Press, New York.

Revelt, D. and K. Train (1998): "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level", *Review of Economics and Statistics* 80, 647-657.

Revelt, D. and K. Train (1999): "Customer-Specific Taste Parameters and Mixed Logit", *Working Paper*, Department of Economics, University of California, Berkeley.

Rizzi, L.I. and J. de D. Ortúzar (2003): "Stated Preference in the Valuation of Interurban Road safety", *Accident Analysis and Prevention* 35, 9-22.

Sawtooth Software (2000): "CBC Hierarchical Bayes Analysis". Sawtooth Software Inc. (http://www.sawtoothsoftware.com/download/techpap/hbtech.pdf)

Sillano, M. and J. de D. Ortúzar (2004): "Willingness-to-Pay Estimation with Mixed Logit Models: Some New Evidence", *Environment and Planning* (in press).

Silva, M. and R.A. Garrido (2003): "Estimation Performance of Low Discrepancy Sequences in Stated Preferences", *Proceedings 10th International Conference on Travel Behaviour Research*, Lucerne, Switzerland, August 2003.

Small, K. and H. Rosen (1981): "Applied Welfare Economics with Discrete Choice Models", *Econometrica* 49, 105-130.

Train, K. (1998): "Recreation Demand Models with Taste Differences Over People", *Land Economics* 74, 230-239.

Train, K. (2001): "A Comparison of Hierarchical Bayes and Maximum Simulated

Likelihood for Mixed Logit", *Working Paper,* Department of Economics, University of California, Berkeley.

Train, K. (2003): *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge.

Von Haefen, R.H. (2003): "Incorporating Observed Choice into the Construction of Welfare Measures from Random Utility Models", *Journal of Environmental Economics and Management* 45, 145-165.

Williams, H.C.W.L. (1977): "On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit", *Environment and Planning* 9A, 285-344.