A discrete choice model incorporating thresholds for perception in attribute values

Víctor Cantillo[a], Benjamin Heydecker[b] and Juan de Dios Ortúzar[c]

Abstract

In this paper we formulate a discrete choice model that incorporates thresholds in the perception of attribute changes. The model considers multiple options and allows changes in several attributes. We postulate that if thresholds exist they could be random, differ between individuals, and even be a function of socio-economic characteristics and choice conditions. Our formulation allows estimation of the parameters of the threshold probability distribution starting from information about choices.

The model is applied to synthetic data and also to real data from a stated preference survey. We found that where perception thresholds exist in the population, the use of models without them leads to errors in estimation and prediction. Clearly, the effect is more relevant when the typical size of change in the attribute value is comparable with the threshold, and when the contribution of this attribute in the utility function is substantial. Finally, we discuss the implications of the threshold model for estimation of the benefits of transport investments, and show that where thresholds exist, models that do not represent them can overestimate benefits substantially.

[a] Department of Civil Engineering, Universidad del Norte. Km 5 Antigua Vía a Puerto Colombia, Barranquilla, Colombia. Tel: 57-5-350 9236; E-mail: vcantill@uninorte.edu.co
[b] Centre for Transport Studies, University College London, Gower Street London WC1E 6BT, England. Tel: 44- 20 7679 1553; E-mail: ben@transport.ucl.ac.uk
[b] Department of Transport Engineering, Pontificia Universidad Católica de Chile, Casilla 306, Código 105, Santiago 22, Chile. Tel: 56-2-686 4822; E-mail: jos@ing.puc.cl

Introduction

Many factors affect the behaviour and decision processes of individuals; these factors are dynamic and change continuously. Moreover, the changes can occur at a macroscopic scale (e.g. social and political processes, environmental quality, urbanization processes), or microscopic scale (e.g. income changes, vehicle purchases). It is important for planners to establish the behaviour of individuals in respect of these changes and therefore the magnitude, direction and timing of their responses. However, individuals seldom acquire information on changes immediately, so this knowledge is imperfect and changes may often be ignored. In addition, in the face of new situations people might undergo a process of experimentation and learning, through trial and error (Kitamura, 1990).

Small changes (absolute or relative) most likely do not prompt any action due to the existence of thresholds in perceptions or *just noticeable differences* (Coombs *et al* 1970). In this sense, perceptible changes would be values above a certain threshold, and those below it would not cause a reaction in the individual (utilities do not change). So, if we define X^t as the value that attribute X takes at time t, the change in utility between t and $t+1$ will only be perceptible to the individual if $|\Delta X^{t+1}|=|X^{t+1}-X^t|>\delta$[1], where δ is a non-negative threshold value.

This phenomenon is complex because changes can accumulate and eventually exceed the threshold, but at the same time there is an adjustment in individual behaviour dependent on the speed of change, that in turn modifies the threshold. Thus, in the most general case the thresholds should be dynamic and depend on the experiences and restrictions of the individual. Therefore, it is possible to postulate that these thresholds are distributed randomly in the population according to the concept of *psychological thresholds* in the theory of consumer choice introduced by Georgescu-Roegem (1958).

To delve into the complexity of the phenomena, according with *Prospect Theory* (Kahneman and Tversky, 1979) the response to changes could be asymmetric; thus, the magnitude of changes in behaviour may be different depending on their direction. In addition, other phenomena such as inertia, habit or reluctance to change may also be present; that is, individual behaviour can be characterized by habit formation making people reluctant to change and consequently the same past behaviour can still prevail because altering it implies time and costs (monetary and psychological). Moreover, change asymmetry can lead to the phenomenon of hysteresis (Goodwin, 1977; Blase, 1979; Williams and Ortúzar, 1982).

Researchers in mathematics, psychology and psychophysics have studied the magnitude of individual response in the presence of changes in the intensity of stimuli in the

---

[1] A more general approach is to express $\Delta X^{t+1} = \left| X^{t+1} - \tilde{X}^{t+1} \right|$, where $\tilde{X}^{t+1}$ is the expected (or reference) value of X^{t+1}. The latter can be defined as an exponentially smoothed function $\tilde{X}^{t+1} = \lambda \tilde{X}^{t} + (1-\lambda) X^{t}$, where λ should be between 0 and 1. We do not search exhaustively on this parameter so we just assume $\lambda=0$. A more suitable value could be obtained using a Bayesian approach or a Kalman Filter (Harrison and Stevens, 1976)

physical world; the existence of thresholds is recognized in this context and they are defined as the minimum size of stimulus required to produce an effect. The relation between the physical stimulus and the response of the mind has been modelled, among others, by Dzhafarov and Colonius (2001). Biostatististicians have proposed models considering dose-response thresholds with random effects (Li and Hunt, 2003), and researchers in Marketing have examined consumer price sensitivity using models that incorporate probabilistic thresholds for price gains and losses with respect to the reference price (Han *et al*, 2001). Finally, Li and Hultkrantz (2000) proposed a model to estimate the value of time considering a stochastic perception threshold; however, the model was restricted to binary choice and assumed a uniformly distributed time threshold and no thresholds for the other attributes.

To summarise, the discussion about thresholds in discrete choice modelling has focused on the duration of time saved in transport projects and on the sensibility to prices in marketing. We intent to extend the discussion to any attribute starting from the hypothesis that it is important to study this concept because if thresholds are not considered, especially when changes are small, it could lead to errors in prediction. Moreover, it is possible to overestimate benefits in transport projects (e.g. time savings) because the impact of each unit of travel time saved below some critical thresholds could amount to nothing that is appreciable for the individual (Welch and Williams, 1997).

In the present paper, we formulate a discrete choice model incorporating thresholds as minimum perceptible changes in attributes. As an important contribution we generalize the model to multiple options and allow changes in several attributes, providing a fairly general approach. We also postulate that if thresholds exist they could be random, differ between individuals and even be a function of socio-economic characteristics and choice conditions. The formulation allows us to estimate the parameters of the threshold probability distribution starting from information about choices. Here, we do not consider either the presence of habit or the asymmetry in thresholds although we recognize that they could be present. Extensions to a non-symmetric treatment are straightforward, but incorporating inertia or habit effects is complicated.

The rest of the paper is organized as follow. In the next section we discuss the theoretical and mathematical formulation of the model including hypotheses, constraints and procedures for its estimation. Then, we present two empirical analyses: one using simulated data and another using data collected as part of a SP survey. In both cases we carry out a comparison of our method with the traditional compensatory model. After that, we consider the implications of the threshold model in the evaluation of transport investment benefits. Finally we present the practical conclusions and implications of our work, and discuss possible avenues for further research.

A threshold model

Consider the case of an individual q who at time t chooses option r and has a set of other available options $A^r(q)$[2]. Let us denote by $\hat{X}_k$ the difference between the values of attribute X_k for options j and r. In addition, let $\boldsymbol{Y}_q$[3] be a vector of socio-economic characteristics of the individual. Then, at time t we have:

$$\hat{X}^t_{kjq} = X^t_{kjq} - X^t_{krq} \tag{1}$$

and assuming a linear in the parameters expression, the utility of option j relative to r can be written as:

$$\hat{U}^t_{jq} = \hat{V}^t_{jq} + \varepsilon^t_{jq} = \boldsymbol{\alpha}\hat{\boldsymbol{X}}^t_{jq} + \boldsymbol{\beta}\boldsymbol{Y}^t_q + \varepsilon^t_{jq} \tag{2}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are vectors of parameters[4]. Then, the probability that individual q chooses option r at time t is:

$$P^t_{rq} = P\left\{\hat{U}^t_{jq} \le \hat{U}^t_{rq} \quad \forall j \in A^r(q)\right\}. \tag{3}$$

Suppose that at time t+1 there are changes in some attributes, for example from $\hat{X}^t_{kjq}$ to $\hat{X}^{t+1}_{kjq}$:

$$\hat{X}^{t+1}_{kjq} = \hat{X}^t_{kjq} + \Delta\hat{X}^{t+1}_{kjq} \tag{4}$$

As an hypothesis we suggest that if the change is small the variation in the individual's systematic utility will be null. Therefore, the individual will only perceive an alteration in his/her utility if $\left|\Delta\hat{X}^{t+1}_{kjq}\right| \ge \delta_{kq}$, where δ_{kq} is a non-negative random perception threshold that is distributed in the population with density function $\phi(\delta_k)$. The individual only perceives that part of $\left|\Delta\hat{X}^{t+1}_{kjq}\right|$ bigger than δ_{kq}, so that the response is to a stimulus of size $\mathrm{Max}\left(\left|\Delta\hat{X}^{t+1}_{kjq}\right| - \delta_{kq},\ 0\right)$[5]. The value of the threshold δ_{kq} can be expressed as a proportion of $\hat{X}^t_{kjq}$ so that:

$$\delta_{kq} = \left|\hat{X}^t_{kjq}\right| \cdot \left(\bar{\delta}_k + \eta_{kq}\right)^6 \tag{5}$$

[2] $A^r(q)$ does not include option r; its size is $||A^r_{(q)}||=J$.
[3] In our notation, bold-face letters represent vectors.
[4] As in this notation $\hat{V}^t_{rq}$ = 0, the utility of r is given only by the error term $\hat{U}^t_{rq} = \varepsilon^t_{rq}$.
[5] An alternative approach is to consider responses to a stimulus of size $\left|\Delta\hat{X}^{t+1}_{kjq}\right|$ whenever this exceeds δ_{kq}.
[6] An alternative approach would be to express the threshold in absolute terms as $\delta_{kq} = \left(\bar{\delta}_k + \eta_{kq}\right)$.

where $\overline{\delta}_k$ is an expected perceived value common to all individuals of similar characteristics and η_{kq} represents individual deviations following a certain probability function with mean zero and variance σ_η^2. If the population is relatively homogenous the means of the distributions may be identical; nevertheless, it is possible to consider variations between individuals by expressing these means as functions of certain attributes of the individuals. For example, a linear in the parameters expression might be used:

$$\overline{\delta}_{kq} = \tau + \boldsymbol{\rho}\mathbf{Y}_q \qquad (6)$$

where τ and $\boldsymbol{\rho}$ are parameters.

As an illustration, Figure 1 shows the variation in utility (ΔV_j) when the attribute X_{jk} (with negative marginal utility) changes *ceteris paribus*. The threshold in this case is 5, so values of ΔX_{jk} between –5 and +5 do not induce changes in utility, whereas outside this interval the utility is affected. Note that here we suppose the threshold to be symmetric.

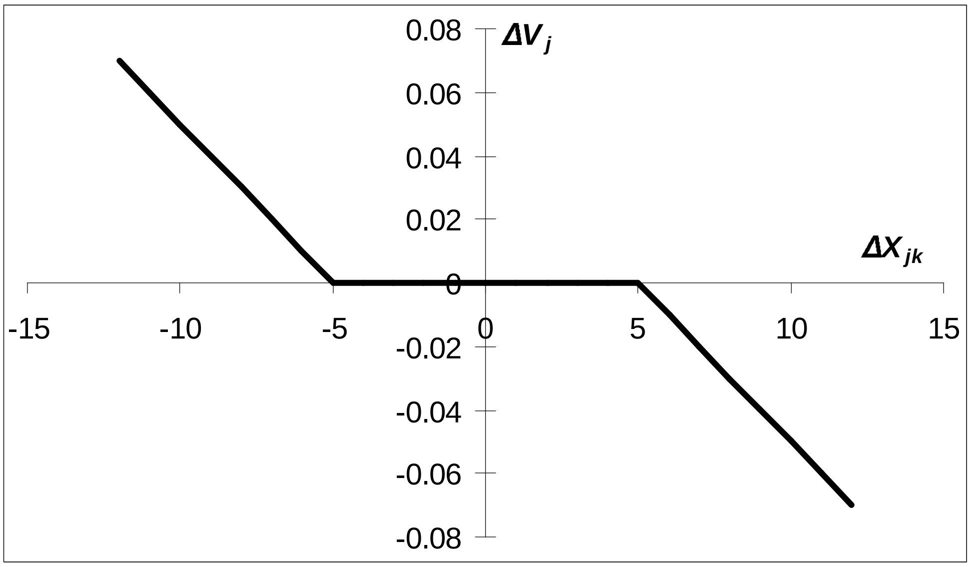

Figure 1. Influence of thresholds in the utility function

After the changes at time t, if we consider thresholds for m of the K attributes the new conditional utility function for option j given the vector of thresholds $\boldsymbol{\delta}_q$ would be:

$$\hat{U}_{jq}^t\left|\boldsymbol{\delta}_q = \hat{V}_{jq}^t\right|\boldsymbol{\delta}_q + \varepsilon_{jq}^t = \sum_{k=1}^{m}\alpha_{kjq}\left[\hat{X}_{kjq}^{t-1} + \left(\Delta\hat{X}_{kjq}^t - \delta_{kq}sign\left(\Delta\hat{X}_{kjq}^t\right)\right)f_{kjq}\right] + \sum_{k=m+1}^{K}\alpha_{kjq}\hat{X}_{kjq}^t + \boldsymbol{\beta}\mathbf{Y}_q^t + \varepsilon_{jq}^t \qquad (7)$$

[7]

where $f_{kjq} = \begin{cases} 1 & if\ \left|\Delta\hat{X}_{kjq}^t\right| \geq \delta_{kq} \\ 0 & \text{otherwise.} \end{cases}$

[7] An alternative and more complicated formulation of the error term is $\varepsilon_{jq}^t = \nu_{jq} + \zeta_j^t + \xi_{jq}^t$, where ν_{jq} is a random term representing an effect that is specific to the individual but invariant over time (that introduces serial correlation), ζ_j^t is a time-specific error affecting all individuals equally and ξ_{jq}^t a purely random error term.

Let the thresholds be distributed in the population according to a join density function $\Omega(\boldsymbol{\delta})$. We will assume that the thresholds are independent between attributes; so we express:

$$\Omega(\boldsymbol{\delta}) = \prod_{k=1}^{m} \phi(\delta_k) \tag{8}$$

The conditional probability that the individual switches from *r* to *j* given the vector of thresholds $\boldsymbol{\delta}$ will be:

$$P_{jq}^{t}\Big|\boldsymbol{\delta} = P\left\{\hat{U}_{jq}^{t}\Big|\boldsymbol{\delta} > \hat{U}_{iq}^{t}\Big|\boldsymbol{\delta}\ \forall i \in A_{(q)}^{r}\ \wedge\ \hat{U}_{jq}^{t}\Big|\boldsymbol{\delta} > \hat{U}_{rq}^{t}\Big|\boldsymbol{\delta}\right\} \tag{9}$$

On the other hand, the probability that the individual does not switch and remains with option *r* given $\boldsymbol{\delta}$, is:

$$P_{rq}^{t}\Big|\boldsymbol{\delta} = P\left\{\hat{U}_{iq}^{t}\Big|\boldsymbol{\delta} \le \hat{U}_{rq}^{t}\Big|\boldsymbol{\delta}\ \forall i \in A_{(q)}^{r}\right\} \tag{10}$$

To calculate the unconditional probability of switching and that of not switching it is necessary to integrate over the values of $\boldsymbol{\delta}$. As an illustration, suppose that random errors are distributed Gumbel IID, in which case we would have the following expression:

$$P_{jq}^{t} = \int_{\mathbf{0}}^{\boldsymbol{\infty}} \frac{\exp\left(V_{jq}^{t}\big|\boldsymbol{\delta}\right)}{\left[1 + \sum_{\forall i \in A_{(q)}^{r}} \exp\left(V_{iq}^{t}\big|\boldsymbol{\delta}\right)\right]} \Omega(\boldsymbol{\delta})\, d\boldsymbol{\delta} \tag{11}$$

In the same vein, the probability that the individual remains using *r* is given by:

$$P_{rq}^{t} = \int_{\mathbf{0}}^{\boldsymbol{\infty}} \frac{1}{\left[1 + \sum_{\forall i \in A_{(q)}^{r}} \exp\left(V_{iq}^{t}\big|\boldsymbol{\delta}\right)\right]} \Omega(\boldsymbol{\delta})\, d\boldsymbol{\delta} \tag{12}$$

Equations (11) and (12) are difficult to solve analytically because they involve multidimensional integrals. Here instead, we employ a probability simulator within a maximum likelihood framework, which leads to Maximum Simulated Likelihood (MSL) estimates.

The log-likelihood function is:

$$L^{t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\delta}) = \sum_{q=1}^{Q}\left[\sum_{j \in A_{(q)}^{r}} g_{jq} \ln\left(P_{jq}^{t}\big|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\delta}\right) + \left(1 - \sum_{j \in A_{(q)}^{r}} g_{jq}\right) \ln\left(P_{rq}^{t}\big|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\delta}\right)\right] \tag{13}$$

where g_{jq} is 1 if the individual switches to *j* and 0 otherwise. The response probability for option *j* is replaced with the unbiased, smooth, tractable simulator (Train, 2003):

$$\hat{P}^t_{jq}\Big|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\delta} = \frac{1}{N}\sum_{n=1}^{N}\Lambda\left(i\,\middle|\,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\delta}^n\right) \tag{14}$$

where $\boldsymbol{\delta}^n$ denotes the *n*th draw from the distribution of $\boldsymbol{\delta}$, and Λ is the probability equation for the logit kernel model specified by equations (11) and (12). In order to generate the set of discrete points we can use pseudo-random sequences (see the discussion in Silva and Garrido, 2003).

Incorporating the simulated probability, the simulated log-likelihood function is then:

$$\hat{L}^t\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\delta}\right) = \sum_{q=1}^{Q}\left[\sum_{J\in A^r_{(q)}} g_{jq}\ln\left(\hat{P}^t_{jq}\middle|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\delta}\right) + \left(1-\sum_{J\in A^r_{(q)}} g_{jq}\right)\ln\left(\hat{P}^t_{rq}\middle|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\delta}\right)\right] \tag{15}$$

It is worth emphasizing that the traditional discrete choice model without thresholds (i.e. in this case a Multinomial Logit, MNL model) is a particular case of the model presented here, where the means and their variances are zero (this implies the absence of any thresholds).

Empirical Analysis

We applied the present threshold model (7) to two data sets; the first is synthetic and the second came from a SP survey. Because the emphasis here is on thresholds rather than on the error term of the utility functions, in all cases we suppose that errors are distributed IID Gumbel.

Application to synthetic data

To examine the performance of the proposed model for a population where thresholds do exist, we followed the Williams and Ortúzar (1982) method of generating a simulated data bank. This consists of three hypothetical options: *Taxi*, *Bus* and *Metro* and includes three attributes: *Cost C*, *Travel Time* T_t and *Access Time* T_a. We used left-truncated Normal distributed attributes and a Normal distributions for relative changes as shown in Table 1.

Table 1. Parameters used for attribute generation

Attribute		Initial Situation			Relative Changes		
		Taxi	Bus	Metro	Case *A*	Case *B*	Case *C*
Cost (*C*)	Mean	45	20	15	0.10	0.25	0.50
	Standard deviation	10	3	1	0.03	0.05	0.07
Travel Time (T_t)	Mean	18	30	10	-0.10	-0.25	-0.50
	Standard deviation	5	10	3	0.03	0.05	0.07
Access Time (T_a)	Mean	8	10	15	-0.10	-0.25	-0.50
	Standard deviation	4	5	5	0.03	0.05	0.07

Thresholds were also generated for each attribute following symmetric triangular distributions, the parameters of which are shown in Table 2. We studied three cases: small, medium and large changes (*A*, *B* and *C* respectively, see Table 2). In the main body of the text we only show case *B*; cases *A* an *C* can be found in the Appendix. We generated 10,000 independent observations according to the following procedure:

1. Initially individuals choose between options according to a compensatory utility maximizing process (i.e. a MNL with scale parameter equal to one). The parameters of the utility function used in the simulation are presented in Table 2.
2. Attributes change and individuals compare these changes with their respective thresholds; if the changes do not exceed the thresholds, the utility function does not change. We assume the triangular threshold distributions start at zero, so they are specified by a single parameter (i.e. mean or mode)
3. After the changes in attributes and the consequent variations in the utility function, the individual chooses again between the options according to the original compensatory utility maximizing process.

Table 2 shows the number of individuals for whom the change in one or other of the options exceeds the respective threshold for each simulated case. To present results in the table, we used the following convention: if an individual chooses *Taxi*, then option 1 will be *Bus* and option 2 *Metro*; is s/he chooses *Bus*, option 1 will be *Taxi* and option 2 *Metro*; finally, if s/he chooses *Metro*, option 1 will be *Taxi* and option 2 *Bus*. Obviously this distribution depends on both the thresholds in Table 2 and the changes in Table 1. It is possible to see that the effect of thresholds is strongest for the attribute *Access Time* and weakest for *Cost.* This is because *Cost* has the smallest marginal utility, and *Access Time* the largest one[8].

Table 2. Parameters used for threshold generation

Parameter	Threshold mean	Value in utility function	Individuals whose threshold is exceeded by an attribute change (Total sample= 10,000)					
			Case A		Case B		Case C	
			Option 1	Option 2	Option 1	Option 2	Option 1	Option 2
Cost	0.08	-0.070	7991	2428	9680	5391	9942	7972
Travel time	0.12	-0.150	3896	5446	6852	8124	8447	9158
Access time	0.16	-0.200	1448	2039	3166	3467	5965	5956

We analysed eight scenarios for each case considering combinations of the presence or absence of thresholds. For instance (see Table 3), the first consists of a population without thresholds in any attribute; by contrast, the last scenario considers a population with thresholds for all attributes. For each scenario, we estimated first a classical MNL model consistent with compensatory behaviour and no thresholds; the second is the proposed model including thresholds for all of the attributes. In parentheses two t-tests appear: the first corresponds to the traditional null hypothesis $\theta_k = 0$, and the second refers to the null hypothesis $\theta_k = \theta_v$, where θ_v are the *true* parameter values, which are referred to as *targets* in the table and are the parameter values in the utility function of Table 2.

Scenario 1 shows that when thresholds do not exist the proposed model converges correctly to the standard MNL, with estimates of threshold parameters close to zero. In general, parameter recovery is successful and the proposed model has always a better fit than the mis-specified MNL model. In fact, in all cases a likelihood ratio (LR)[9] test

[8] The latter is a consequence of having based the simulated data on real data and models estimated previously.

[9] $LR=-2\{l(\theta_r)- l(\theta)\}$, where $l(\theta_r)$ is the log-likelihood at convergence for a restricted version of a more general model with log-likelihood $l(\theta)$. LR is distributed asymptotically χ^2 with r degrees of freedom; r is

shows that the proposed model is preferable at the 5% level to the MNL. In addition, almost always the null hypothesis $\theta_k = \theta_v$ is accepted at the same level, and only in few cases the *t*-tests show that the estimated thresholds parameters are different from their respective targets. Furthermore, the parameters in the utility function are recovered accurately in all cases.

It is worth noting that the MNL does not generally give adverse diagnostics in this case (i.e. in 3 of 7 scenarios the parameters in the utility function are well recovered). Simulations show that when changes in the attributes increase the difference in fit between the proposed model and the MNL increases too, but the accuracy in parameter recovery decreases. On the other hand, as the incidence of the *Access Time* threshold is stronger than the others, its estimation tends to be more successful.

In order to investigate the consequences of adopting models with and without thresholds when they do exist, we tested the performance of the MNL and of the proposed model estimated for scenario 8 in case B (Table 3), in terms of their predictive capabilities. Consequently, we used plans representing “policy changes” which ranged from slightly to substantially different from the base data used for estimation. This entailed changing attribute values for the options and re-executing the choice simulation procedure. We generated a series of simulated future scenarios, which could be compared with the model predictions.

Following the approach of Munizaga *et al* (2000) we tested six policies *P*1 to *P*6; the specific changes in attributes associated to each one are shown in Table 4. Policies *P*1 to *P*3 correspond to small changes; in contrast, *P*4 to *P*6 represent aggressive policy changes. The error measure considered was the percentage difference between the behaviour that was simulated for the modified attribute values and that estimated with the model under scrutiny. The minimum response error that may be considered a prediction error is given by the standard deviation of the simulated observations generated with different seeds, as it reflects the inherent variability of the simulation process.

We carried out 15 repetitions of the data generation process for each policy using different seeds for the pseudo-random numbers and found that the largest value of the coefficient of variation was 8.3%. Given these results, we decided to take 10% as a reasonable tolerance error. Therefore, any discrepancy exceeding this value was considered an estimation error.

---

the number of linear restrictions needed to pass from the non restricted threshold model to the MNL (Ortúzar and Willumsen, 2001, p263)

Table 3. Models with simulated database, Case *B*, medium size changes; n = 10,000

Threshold		Scenario															
		1		2		3		4		5		6		7		8	
Cost				√						√		√				√	
Travel Time						√				√				√		√	
Access Time								√				√		√		√	
Individuals changing their choice when threshold present		0		142		251		316		179		247		544		427	
Model	Target	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed
Cost	-0.070	-0.071 (-48.8) [-0.6]	-0.071 (-44.3) [-0.8]	-0.069 (-48.0) [0.6]	-0.071 (-48.0) [-0.3]	-0.072 (-47.8) [-1.4]	-0.070 (-46.5) [-0.3]	-0.073 (-47.1) [-1.6]	-0.071 (-46.8) [-0.6]	-0.071 (-47.7) [-0.6]	-0.070 (-46.3) [-0.1]	-0.072 (-47.1) [-1.1]	-0.071 (-46.5) [-0.5]	-0.074 (-46.8) [-2.3]	-0.071 (-45.2) [-0.3]	-0.073 (-46.9) [-2.2]	-0.071 (-44.9) [-0.6]
Travel Time	-0.150	-0.147 (-53.0) [1.3]	-0.147 (-51.2) [1.1]	-0.145 (-53.1) [2.0]	-0.147 (-53.6) [1.2]	-0.151 (-53.1) [-0.4]	-0.148 (-51.2) [0.8]	-0.149 (-52.5) [0.3]	-0.146 (-52.1) [1.4]	-0.150 (-53.1) [-0.1]	-0.149 (-52.0) [0.3]	-0.147 (-52.4) [1.1]	-0.145 (-52.2) [1.7]	-0.152 (-53.3) [-0.6]	-0.147 (-51.2) [1.2]	-0.150 (-52.3) [-0.1]	-0.147 (-49.1) [1.1]
Access Time	-0.200	-0.198 (-39.1) [0.4]	-0.201 (-39.0) [-0.2]	-0.198 (-39.5) [0.5]	-0.199 (-39.7) [0.1]	-0.197 (-38.3) [0.6]	-0.197 (-38.1) [0.7]	-0.181 (-35.5) [3.8]	-0.196 (-35.4) [0.7]	-0.196 (-38.4) [0.8]	-0.197 (-38.4) [0.6]	-0.179 (-35.4) [4.2]	-0.193 (-30.3) [1.1]	-0.177 (-34.4) [4.6]	-0.193 (-34.2) [1.3]	-0.178 (-34.7) [4.4]	-0.193 (-34.0) [1.2]
Mean Threshold *Cost*	0.080	-	0.002 (0.2) [0.2]	-	0.084 (9.0) [0.5]	-	-	-	-	-	0.071 (4.4) [-0.5]	-	0.052 (4.5) [-2.4]	-	-	-	0.059 (4.1) [-1.4]
Mean Threshold *Travel Time*	0.120	-	0.000 (0.0) [0.0]	-	-	-	0.105 (7.0) [-1.0]	-	-	-	0.094 (6.6) [-1.8]	-	-	-	0.101 (9.8) [-1.9]	-	0.090 (4.1) [-1.4]
Mean Threshold *Access Time*	0.160	-	0.008 (1.2) [1.2]	-	-	-	-	-	0.149 (13.7) [-1.0]	-	-	-	0.128 (3.0) [-0.8]	-	0.162 (10.6) [0.1]	-	0.150 (10.6) [-0.7]
Log-likelihood		-5751.8	-5751.7	-5859.5	-5851.8	-5595.3	-5577.9	-5512.8	-5489.9	-5654.4	-5646.7	-5581.4	-5570.8	-5380.9	-5314.1	-5427.2	-5391.3
LR		0.2		15.3		34.9		45.8		15.5		21.3		133.6		71.7	
Critical χ^2 at 5% level		7.81		3.84		3.84		3.84		5.99		5.99		5.99		7.81	

Table 4. Policy changes: percentage change in attribute values

Policy	Cost Bus	Cost Metro	Travel Time Bus	Travel Time Metro	Access Time Taxi	Access Time Bus	Access Time Metro
P1			-20				
P2		-10					
P3					-15		
P4			-50				
P5		100					
P6	-50	50	-50	50		-50	50

To test goodness of fit we used the following Chi-squared measure:

$$\chi^2 = \sum_i \frac{\left(\hat{N}_i - N_i\right)^2}{N_i}, \tag{16}$$

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 result should be compared with the critical χ^2 value at the 5% level with two degrees of freedom, which is 5.99.

Table 5 shows that the proposed model always provides superior results to the mis-specified MNL, with response errors well within the tolerance of 10%. By contrast, the MNL yields response errors larger than this value in several cases (*P*1, *P*4 and *P*6). Furthermore, the χ^2 index shows that the errors for the proposed model are not statistically significant at the tested level but the model errors for the MNL are significant in the above cases. These results allow us to infer that if there are thresholds for perceptible minimum changes in attributes, use of a mis-specified no-threshold model can lead to prediction errors.

Table 5. Comparison of simulated and modelled forecasts

Policy	Target			MNL				Proposed Model			
	Taxi	Bus	Metro	Taxi	Bus	Metro	χ^2	Taxi	Bus	Metro	χ^2
*P*1	108	147	745	87	164	749	6.1	100	155	745	1.0
				-19.4%	11.6%	0.5%		-7.4%	5.4%	0.0%	
*P*2	110	110	780	102	109	789	0.7	108	112	780	0.1
				-7.3%	-0.9%	1.2%		-1.8%	1.8%	0.0%	
*P*3	112	108	780	109	101	790	0.7	108	109	783	0.2
				-2.7%	-6.5%	1.3%		-3.6%	0.9%	0.4%	
*P*4	84	308	608	68	330	602	4.7	79	312	609	0.4
				-19.0%	7.1%	-1.0%		-6.0%	1.3%	0.2%	
*P*5	195	170	635	184	180	636	1.2	189	179	632	0.7
				-5.6%	5.9%	0.2%		-3.1%	5.3%	-0.5%	
*P*6	63	832	105	52	874	74	13.2	60	839	101	0.4
				-17.5%	5.0%	-29.5%		-4.8%	0.8%	-3.8%	

To evaluate the effect of sample size on model estimation we generated simulated data with only 1,000 draws. Similar to before, case B is presented in Table 6 whilst the others can be found in the Appendix. In these cases, the recovery of the true parameters is similar to the previous experiment but the improvements in log-likelihood are less pronounced, being significantly different from zero in 4 out of the 7 cases simulated with thresholds. In fact, excluding scenarios 4, 7 and 8, the χ^2 tests show that the proposed model and the MNL are equivalent at the 5% level. This confirms the requirement for adequate sample size in estimation, especially when dealing with complex models including many attributes (Williams and Ortúzar, 1982; Munizaga *et al*, 2000).

The results of the simulation allow us to conclude that the threshold parameters are reasonably well recovered if the sample size is large enough, especially when the effect of the attribute is strong (high marginal utility and/or high threshold mean). In these cases the model without thresholds may lead to errors in estimation and in prediction. Conversely, when the effect of the attribute is weak (low marginal utility and/or threshold mean close to zero), or if the sample size is not large enough, the threshold parameters are not so well recovered and, also, in these cases the errors using the traditional MNL are small.

Application to a stated preference survey

The data for this part of the analysis consists of a route choice SP survey for car-trips (Caussade *et al*, 2004). We selected part of the survey with choices based on three attributes: *Travel time* (min), *Trip time variability* (min), and *Total cost* (US$). In the implementation of the experiment, respondents were first asked to consider a trip they had taken recently and to report its attributes (this was called *Current Route*). Then a computer program automatically generated the hypothetical choice scenarios according to a fractional factorial design. Each specific design pivoted on the attribute levels associated with the *Current Route*. As a generic design, the added options (two in the case of this data) were of exactly the same nature.

The total number of valid observations for our analysis was 718. Of these, 247 (34.4%) chose the current route. On the other hand, there were 108 (16%) respondents who answered lexicographically; of these, 96 were lexicographic in the *Travel time* variable and 12 in the *Total cost* variable. No observation was found to be lexicographic on *Trip time variability*.

Model TM1 in Table 7 includes a threshold for *Travel Time,* model TM2 incorporates a threshold for *Variability* and model TM3 a threshold for *Cost*. In all cases we assumed symmetric Triangular distributions for the thresholds. As seen, models TM2 and TM3 are equivalent to the MNL, with estimated threshold means close to 0 ($< 10^{-4}$) and unchanged log-likelihood; this means that there is no evidence of thresholds for the attributes *Variability* and *Cost*. In marked contrast, model TM1 is significantly better than the reference MNL model (LR= 33.8 is substantially higher than 3.84, the critical value at the 5% level); consequently, we can conclude that a threshold for *Travel Time* exists with a mean close to 12% of its initial value. Note that its presence results in a substantially increased magnitude for the coefficient of travel time (i.e. from -0.0353 to -0.1299); it also results in an increase in the magnitude of the cost coefficient (i.e. from -0.2100 to -0.420).

Table 6. Models with simulated database, Case B, medium size changes; *n* = 1,000

Threshold		Scenario															
		1		2		3		4		5		6		7		8	
Cost				√						√		√				√	
Travel Time						√				√				√		√	
Access Time								√				√		√		√	
Individuals changing their choice when threshold present		0		18		26		34		21		32		51		46	
Model	Target	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed
Cost	-0.070	-0.073 (-15.6) [-0.7]	-0.073 (-14.7) [-0.7]	-0.072 (-15.7) [-0.3]	-0.072 (-15.6) [-0.4]	-0.072 (-15.3) [-0.5]	-0.071 (-15.0) [-0.2]	-0.073 (-15.0) [-0.7]	-0.072 (-15.0) [-0.5]	-0.069 (-15.2) [0.2]	-0.068 (-14.9) [0.4]	-0.071 (-15.0) [-0.1]	-0.071 (-14.7) [-0.2]	-0.072 (-14.9) [-0.4]	-0.070 (-14.4) [0.1]	-0.073 (-15.0) [-0.6]	-0.071 (-14.3) [-0.1]
Travel Time	-0.150	-0.142 (-15.5) [0.9]	-0.142 (-15.8) [0.9]	-0.140 (-16.5) [1.2]	-0.140 (-16.1) [1.2]	-0.147 (-16.6) [0.3]	-0.145 (-16.5) [0.5]	-0.149 (-16.4) [0.1]	-0.147 (-16.3) [0.4]	-0.146 (-16.7) [0.5]	-0.144 (-15.9) [0.7]	-0.146 (-16.5) [0.4]	-0.147 (-15.0) [0.3]	-0.154 (-16.5) [-0.4]	-0.150 (-16.2) [0.0]	-0.150 (-16.4) [0.0]	-0.146 (-15.2) [0.4]
Access Time	-0.200	-0.201 (-12.5) [-0.1]	-0.201 (-12.4) [-0.1]	-0.205 (-12.8) [-0.3]	-0.205 (-12.8) [-0.3]	-0.194 (-11.9) [0.4]	-0.195 (-11.9) [0.3]	-0.177 (-10.8) [1.4]	-0.198 (-10.9) [-0.1]	-0.189 (-11.8) [0.7]	-0.189 (-11.8) [0.7]	-0.172 (-10.7) [1.8]	-0.191 (-10.5) [0.5]	-0.177 (-10.7) [1.4]	-0.199 (-10.8) [0.1]	-0.177 (-10.8) [1.4]	-0.196 (-10.3) [0.2]
Mean Threshold *Cost*	0.08	-	0.000 (0.0) [0.0]	-	0.012 (0.2) [-1.2]	-	-	-	-	-	0.000 (0.0) [-1.2]	-	0.025 (0.2) [-0.5]	-	-	-	0,010 (0.1) [-0.7]
Mean Threshold *Travel Time*	0.12	-	0.000 (0.0) [0.0]	-	-	-	0.097 (5.1) [-1.2]	-	-	-	0.073 (3.2) [-2.1]	-	-	-	0.081 (2.1) [-1.0]	-	0,077 (0.6) [-0.3]
Mean Threshold *Access Time*	0.16	-	0.000 (0.0) [0.0]	-	-	-	-	-	0.201 (5.6) [1.1]	-	-	-	0.108 (4.0) [-1.9]	-	0.250 (3.3) [1.2]	-	0,177 (1.6) [0.2]
Log-likelihood		-568.3	-568.3	-582.5	-582.4	-555.0	-552.7	-534.4	-530.2	-567.5	-566.2	-548.5	-546,1	-529.4	-520.5	-535.1	-528,3
LR		0.0		0.2		4.7		8.3		2.5		4.8		17.8		13.6	
Critical χ^2 at 5% level		7.81		3.84		3.84		3.84		5.99		5.99		5.99		7.81	

Table 7. Models for the SP survey

Parameter	MNL	TM1	TM2	TM3
Travel Time	-0.0353	-0.1299	-0.0357	-0.0355
	(-14.8)	(-7.3)	(-14.8)	(-14.8)
Variability	-0.0146	-0.0156	-0.0149	-0.0147
	(-3.4)	(-3.0)	(-3.1)	(-3.5)
Cost	-0.2100	-0.420	-0.2100	-0.2100
	(-8.7)	(-8. 6)	(-8.7)	(-8.1)
Threshold *Travel Time* mean		0.1247		
		(17.0)		
Threshold *Variability* mean			0.0000	
			(0.00)	
Threshold *Cost* mean				0.000
				(0.00)
Number of Observations	782	782	782	782
Log-likelihood	-623.2	606.3	623.2	623.2
LR		33.8	0.0	0.0

Implications for the Evaluation of Transport Investments

Apart from forecasting the impacts of transport policy on demand, random utility models are widely used to estimate Willingness-To-Pay (WTP) for improvements in attributes. For example, the subjective value of time (SVT) is defined as the marginal rate of substitution between time and money at constant utility (Gaudry *et al*, 1989). If the utility function is linear in the parameters and there are no thresholds, the SVT is computed as the ratio of the time and cost parameters ($SVT=\alpha_t/\alpha_c$).

However, in the presence of thresholds as defined in this paper the estimation of benefits is more complicated. For example, if the travel time saved is too small the individual could not perceive it and therefore the benefit of each unit of travel time saved would be reduced (possibly to zero). From the utility function (7) we can derive the expected value (compensated variation) per unit of time; therefore, the conditional SVT in the proposed threshold model is given by:

$$SVT\,|\boldsymbol{\delta} = E\left(-\frac{\Delta c}{\Delta t}|\boldsymbol{\delta}\right) = \left[\frac{\alpha_t f_t}{\alpha_c f_c}\left(1-\frac{\delta_t sign(\Delta t)}{\Delta t}\right)-\frac{\delta_c sign(\Delta c)}{\Delta c}\right] \tag{17}$$

and the unconditional *SVT* would be:

$$SVT = E\left(-\frac{\Delta c}{\Delta t}\right) = \iint\left[\frac{\alpha_t f_t}{\alpha_c f_c}\left(1-\frac{\delta_t sign(\Delta t)}{\Delta t}\right)-\frac{\delta_c sign(\Delta c)}{\Delta c}\right]\phi(\delta_t)\phi(\delta_c)\,d\delta_t d\delta_c \tag{18}$$

Expressions (17) and (18) cannot be used if $f_c = 0$; they only are valid when $\Delta c \geq \delta c$. In order to compute (18), it is necessary to solve the integral; for this, Monte Carlo simulation can be used. Note that the SVT depends on the values of Δc and Δt. If there is no evidence of a threshold for cost, the computation is easier because we have a one-dimensional integral.

On the other hand, if errors are IID Gumbel and the indirect utility function *U* has a common linear income effect we can calculate the conditional *WTP* as follows (McFadden, 1998):

$$E\left[WTP\middle|\boldsymbol{\delta}\right]=\frac{1}{\alpha_c}\left\{Ln\sum_{A^r{}_{(q)}\cup A_r}\exp\left(V_{\hat{i}}^{t}(\boldsymbol{\delta})\right)-Ln\sum_{A^r{}_{(q)}\cup A_r}\exp\left(V_{\hat{i}}^{t-1}\right)\right\} \tag{19}$$

where the indirect utility functions are defined in (2) and (7). Because the vector of thresholds δ is in fact not known, the unconditional WTP is:

$$E\left[WTP\right]=\int_{\boldsymbol{\delta}}\frac{1}{\alpha_c}\left\{Ln\sum_{A^r{}_{(q)}\cup A_r}\exp\left(V_{\hat{i}}^{t}(\boldsymbol{\delta})\right)-Ln\sum_{A^r{}_{(q)}\cup A_r}\exp\left(V_{\hat{i}}^{t-1}\right)\right\}\Omega(\boldsymbol{\delta})\,d\boldsymbol{\delta} \tag{20}$$

Thus, we can estimate (20) by Monte Carlo simulation as follows:

$$\hat{E}\left[WTP\right]=\frac{1}{N}\sum_{n=1}^{N}E\left[WTP\middle|\boldsymbol{\delta}^n\right] \tag{21}$$

where $\boldsymbol{\delta}^n$ denotes the *n*th draw from the distribution of $\boldsymbol{\delta}$.

To investigate the effect of using the threshold model in the evaluation of benefits due to improvements in a transport system, we applied it to a test example using the SP database and compared the results with those of the traditional non threshold model (MNL). We considered three hypothetical scenarios corresponding to savings of 10, 20 and 30 minutes in travel time for each of the 782 individuals of the sample. It is necessary to point out that the range of travel times in the sample is 55 to 600 minutes, with mean and standard deviation close to 230 and 160 minutes respectively. By comparison, the expected value of the threshold for travel time is close to 30 minutes.

For the MNL model, the SVT is a constant given by the ratio between the parameters of travel time and cost: 0.1681 US$/min. In the case of the threshold model TM1, however, the SVT is not constant across individuals, but depends on how large the saving in travel time is in comparison with the current travel time value (see Table 8).

Table 8. Benefit estimation by MNL and proposed threshold models

	Travel time reduction (minutes)		
Estimated Benefit (US$)	10	20	30
Threshold Model	186	1169	2796
MNL Model	1315	2629	3944
Difference (%)	607.0	124.8	41.0
Average SVT (US$/min)			
Threshold Model	0.0238	0.0748	0.1192
MNL Model	0.1681	0.1681	0.1681

As can be seen, the MNL overestimates the benefits although the proportionate difference between the approaches decreases as the saving in travel time increases. As can be seen, when the saving in travel time is 10 minutes, the benefits estimated with the MNL are more than six times those estimated with the threshold model, but when the saving in travel time is 30 minutes, the difference decreases to 41%.

Figure 2 shows the variation in SVT depending on the current travel time and the travel time saved. As we noted, the estimation with the MNL is invariant with respect to these variables but the threshold model is sensitive to them, especially when they are small. The upper limit for SVT in the threshold model is given by the ratio between the parameters of travel time and cost (i.e. US$ 0.3093/min), close to twice the value estimated by the MNL model.

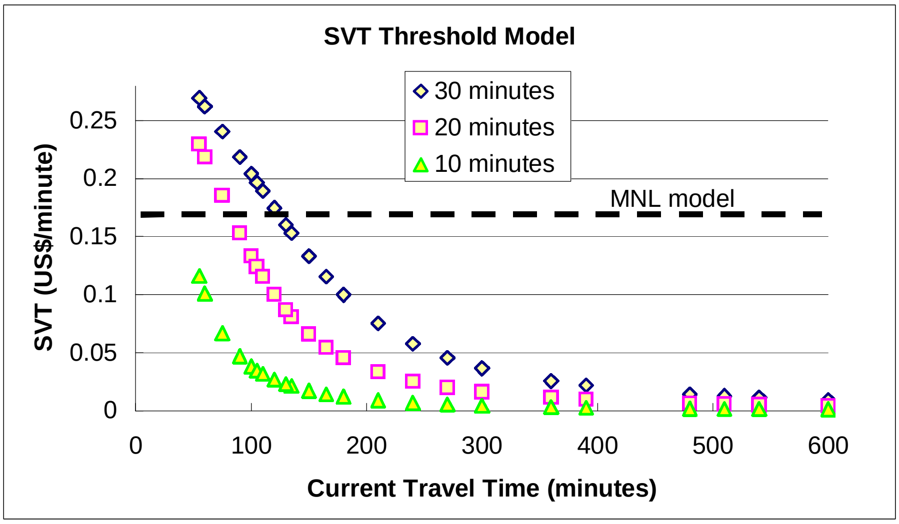

Fig 2. Variation of SVT in the Threshold Model

Conclusions

We propose a discrete choice model that incorporates random thresholds as minimum perceptible changes in attributes, multiple options and changes in several attributes. Our formulation allows for estimation of the parameters of the threshold probability distribution starting from information about choices. The model is of sufficient complexity that its calibration requires use of techniques such as simulated maximum likelihood.

The model was applied to synthetic data and then to real data collected as part of a SP survey. We found that where perception thresholds exist in the population, the use of models without them leads to errors in estimation and in prediction, although this occurs with more emphases when the contribution of a given attribute in the utility function is strong. On the other hand, threshold effects for several attributes can be confounded; for this reason, it could be convenient to test for thresholds individually.

Investigation of data from a stated preference experiment shows that there is evidence for the existence of thresholds in the travel time variable, with an estimated mean size of about 12% of the initial value (note that we considered symmetric thresholds). On the other hand, estimates of thresholds for the variability of travel time and for cost were both close to zero.

The estimation of benefit measures using the threshold model depends on the size of the change in a variable (in the present example, travel time) by comparison with its current value. Moreover, the use of a model without thresholds such as the MNL could substantially overestimate benefits in cases where thresholds exist in the variable that is being changed, although the difference decreases as the magnitude of the change increase.

There are several aspects of interest that remain for future research. One is to evaluate the impact of sample size on the estimation of models including thresholds; a first analysis suggests that the influence of the sample size is strong, especially in terms of improvements in log-likelihood. Another aspect relates to a more in-depth analysis on the application of the threshold model in the evaluation of transport investment benefits to real projects. Finally, the presence of asymmetries in the thresholds is also worthy of investigation.

Acknowledgements

We are grateful to Luis Ignacio Rizzi and Huw Williams for their intellectual help in developing these ideas. Thanks are also due to the Chilean Fund for the Development of Scientific and Technological Research (FONDECYT) for having provided support for this investigation through several projects (1970117, 1000616 and 1020981). Finally, we wish to acknowledge the support of the School of Engineering at the Pontificia Universidad Católica de Chile and the MECE program of the Chilean government for having partially supported a visit of the first author to the Centre of Transport Studies at University College London where the first draft of this paper was written.

References

BLASE, J.H. (1979). Hysteresis and catastrophe theory: a demonstration of habit and threshold effects in travel behaviour. *Proceedings 7th PTRC Summer Annual Meeting*, University of Warwick, July 1979, England.

CAUSSADE, S., ORTUZAR, J. de D., RIZZI, L.I. and HENSHER, D.A. (2004). Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research* 40B (in press).

COOMBS, C.H., DAWES, R.M. and TVERSKY, A. (1970). *Mathematical Psychology*: *An Elementary Introduction*. Prentice-Hall, Englewood Cliffs, NJ.

DZHAFAROV, E.N. and COLONIUS, H. (2001). Multidimensional Fechnerian scaling: basics. *Journal of Mathematical Psychology* 45, 670-919.

GAUDRY, M.J.I., JARA-DIAZ, S.R. and ORTUZAR, J. de D. (1989). Value of time sensitivity to model specification. *Transportation Research* 23B, 151-158.

GEORGESCU-ROEGEN, N. (1958). Thresholds in choice and the theory of demand. *Econometrica* 26, 157-168.

GOODWIN, P. (1977). Habit and hysteresis in mode choice. *Urban Studies* 14, 95-98.

HAN, S., GUPTA, S. and LEHMANN, D. (2001). Consumer price sensitivity and price thresholds. *Journal of Retailing* 77, 435-456.

HARRISON, P.J. and STEVENS, C.F. (1976). Bayesian forecasting. *Journal of the Royal Statistical Society* 38B, 205-247.

KAHNEMAN, D. and TVERSKY, A. (1979). Prospect theory: an analysis of decision making under risk. *Econometrica* 47, 263-291.

KITAMURA, R. (1990). Panel analysis in transportation planning: an overview. *Transportation Research* 24A, 401-415.

LI, Ch-Z. and HULTKRANTZ, L.A. (2000). Stochastic threshold model for estimating the value of travel time. *EAERE 2000 Annual Conference*, Rethymnon, Greece.

LI, Ch-S. and HUNT, D. (2003). Regression splines for threshold selection with application to a random-effects logistic dose-response model. *Computational Statistics and Data Analysis* (in press).

McFADDEN, D. (1998). Measuring willingness-to-pay for transportation improvements. In T. Gärling, T. Laitila, and K. Westin (eds.), *Theoretical Foundations of Travel Choice Modelling*, 339-364, Elsevier Science, Amsterdam.

MUNIZAGA, M.A., HEYDECKER, B.G. and ORTUZAR, J. de D. (2000) Representation of heteroscedasticity in discrete choice models. *Transportation Research* 34B, 219-240.

ORTUZAR, J. de D. and WILLUMSEN, L.G. (2001) *Modelling Transport*. Third Edition, John Wiley and Sons, Chichester.

SILVA, M. and GARRIDO, R.A. (2003) Estimation performance of low discrepancy sequences in stated preference. *Proceedings 10th International Conference on Travel Behaviour Research*, Lucerne, Switzerland.

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

WELCH, M. and WILLIAMS, H.C.W.L. (1997). The sensitivity of transport investment benefits to the evaluation of small time savings. *Journal of Transport Economics and Policy* 31, 231-254.

WILLIAMS, H.C.W.L. and ORTUZAR, J. de D. (1982) Behavioural theories of dispersion and the mis-specification of travel demand models. *Transportation Research* 16B, 167-219.

APPENDIX

Table A1. Models with simulated database, Case A , small changes; n = 10,000

Threshold		Scenario															
		1		2		3		4		5		6		7		8	
Cost				√						√		√				√	
Travel Time						√				√				√		√	
Access Time								√				√		√		√	
Individuals changing their choice when threshold present		0		105		172		196		136		152		365		292	
Model	Target	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed
Cost	-0.070	-0.070 (-45.3) [0.3]	0.069 (-45.0) [0.3]	-0.068 (-45.1) [1.1]	-0.070 (-44.4) [-0.3]	-0.072 (-45.3) [-1.1]	-0.070 (-44.2) [0.1]	-0.071 (-45.2) [-0.9]	-0.070 (-45.0) [-0.1]	-0.070 (-45.2) [-0.2]	-0.070 (-43.4) [0.3]	-0.070 (-45.0) [-0.1]	-0.070 (-43.3) [0.2]	-0.074 (-45.2) [-2.3]	-0.071 (-44.0) [-0.4]	-0.073 (-45.1) [-1.6]	-0.070 (-34.5) [0.1]
Travel Time	-0.150	-0.146 (-53.2) [1.4]	-0.146 (-53.1) [1.4]	-0.145 (-53.3) [1.9]	-0.147 (-53.6) [1.2]	-0.151 (-53.1) [-0.4]	-0.146 (-50.50) [1.4]	-0.149 (-52.9) [0.2]	-0.146 (-52.5) [0.8]	-0.150 (-53.2) [0.0]	-0.147 (-52.7) [1.0]	-0.147 (-53.0) [0.9]	-0.146 (-52.9) [1.6]	-0.155 (-52.8) [-1.7]	-0.148 (-51.7) [0.9]	-0.153 (-52.8) [-1.0]	-0.147 (-36.8) [0.8]
Access Time	-0.200	-0.195 (-35.5) [1.0]	-0.195 (-35.5) [1.0]	-0.193 (-35.5) [1.4]	-0.195 (-35.8) [0.8]	-0.198 (-35.3) [0.4]	-0.195 (-35.0) [1.0]	-0.190 (-34.3) [1.80]	-0.194 (-34.4) [1.1]	-0.196 (-35.4) [0.7]	-0.196 (-35.2) [0.8]	-0.188 (-34.3) [2.1]	-0.194 (-34.4) [1.1]	-0.195 (-34.3) [0.9]	-0.196 (-34.1) [0.7]	-0.193 (-34.3) [1.3]	-0.194 (-25.0) [0.8]
Mean Threshold *Cost*	0.080	-	0.000 (0.0) [0.0]	-	0.102 (7.2) [1.5]	-	-	-	-	-	0.060 (4.0) [-1.3]	-	0.050 (2.7) [-1.4]	-	-	-	0.020 (0.2) [-0.9]
Mean Threshold *Travel Time*	0.120	-	0.000 (-0.8) [-0.8]	-	-	-	0.142 (4.7) [0.7]	-	-	-	0.114 (9.2) [-0.5]	-	-	-	0.120 (8.3) [0.0]	-	0.097 (4.0) [0.9]
Mean Threshold *Access Time*	0.160	-	0.000 (0.0) [0.0]	-	-	-	-	-	0.171 (3.0) [0.2]	-	-	-	0.175 (7.5) [0.6]	-	0.176 (5.4) [0.5]	-	0.143 (8.7) [-1.0]
Log-likelihood		-5458.4	-5458.4	-5532.7	-5529.6	-5287.3	-5276.3	-5292.0	-5278.8	-5358.9	-5350.8	-5367.5	-5359.3	-5104.1	-5066.0	-5180.0	-5155.9
LR		0.0		6.3		22.0		26.2		16.0		16.3		76.1		48.2	
Critical χ^2 at 5% level		7.81		3.84		3.84		3.84		5.99		5.99		5.99		7.81	

Table A2. Models with simulated database, Case C, strong changes; n = 10,000

Threshold		Scenario															
		1		2		3		4		5		6		7		8	
Cost				√						√		√				√	
Travel Time						√				√				√		√	
Access Time								√				√		√		√	
Individuals changing their choice when threshold present		0		169		308		461		200		393		779		629	
Model	Target	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed
Cost	-0.070	-0.070 (-52.5) [-0.1]	-0.070 (-50.4) [-0.2]	-0.069 (-52.6) [0.5]	-0.071 (-51.5) [-0.5]	-0.073 (-52.2) [-1.8]	-0.071 (-50.7) [-0.8]	-0.072 (-51.1) [-1.4]	-0.070 (-50.1) [0.1]	-0.071 (-52.4) [-1.0]	-0.071 (-51.3) [-0.5]	-0.072 (-51.3) [-1.1]	-0.071 (-49.3) [-0.5]	-0.074 (-50.6) [-2.8]	-0.071 (-48.5) [-0.5]	-0.074 (-50.8) [-2.5]	-0.071 (-48.9) [-0.7]
Travel Time	-0.150	-0.146 (-49.3) [1.5]	-0.146 (-46.1) [1.2]	-0.143 (-49.1) [2.4]	-0.146 (-49.3) [1.5]	-0.147 (-49.1) [1.1]	-0.145 (-48.6) [1.6]	-0.148 (-48.9) [0.6]	-0.143 (-48.4) [2.3]	-0.145 (-49.1) [1.6]	-0.145 (-49.1) [1.8]	-0.147 (-48.9) [1.0]	-0.144 (-48.4) [2.1]	-0.151 (-48.8) [-0.3]	-0.146 (-48.5) [1.4]	-0.150 (-48.8) [0.0]	-0.146 (-48.5) [1.5]
Access Time	-0.200	-0.198 (-45.8) [0.5]	-0.198 (-44.2) [0.5]	-0.199 (-46.3) [0.2]	-0.198 (-46.0) [0.5]	-0.193 (-44.5) [1.7]	-0.196 (-44.4) [0.9]	-0.174 (-40.9) [6.2]	-0.195 (-34.6) [0.8]	-0.195 (-45.1) [1.3]	-0.197 (-45.1) [0.8]	-0.175 (-41.3) [5.9]	-0.198 (-37.0) [0.4]	-0.168 (-39.3) [7.6]	-0.200 (-38.0) [0.1]	-0.170 (-39.9) [7.0]	-0.197 (-37.0) [0.5]
Mean Threshold *Cost*	0.080	-	0.013 (0.3) [0.3]	-	0.107 (4.4) [1.1]	-	-	-	-	-	0.060 (3.8) [-1.2]	-	0.107 (4.5) [1.1]	-	-	-	0.071 (4.0) [-0.5]
Mean Threshold *Travel Time*	0.120	-	0.000 (0.0) [0.0]	-	-	-	0.092 (5.2) [-1.6]	-	-	-	0.085 (6.9) [-2.9]	-	-	-	0.084 (7.0) [-3.0]	-	0.076 (7.4) [-4.3]
Mean Threshold *Access Time*	0.160	-	0.000 (0.0) [0.0]	-	-	-	-	-	0.170 (7.4) [0.4]	-	-	-	0.190 (10.8) [1.7]	-	0.205 (15.6) [3.4]	-	0.192 (12.0) [2.0]
Log-likelihood		-6367.1	-6366.8	-6445.8	-6433.3	-6241.0	-6225.4	-6164.3	-6134.4	-6318.6	-6314.2	-6207.7	-6189.1	-6007.5	-5914.1	-6056.2	-6006.4
LR		0.4		25.0		31.1		59.9		8.8		37.2		186.9		99.7	
Critical χ^2 at 95% level		7.81		3.84		3.84		3.84		5.99		5.99		5.99		7.81	

Table A3. Models with simulated database, Case A, small changes; *n* = 1,000

Threshold		Scenario															
		1		2		3		4		5		6		7		8	
Cost				√						√		√				√	
Travel Time						√				√				√		√	
Access Time								√				√		√		√	
Individuals changing their choice when threshold present		0		9		19		23		10		18		34		28	
Model	Target	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed
Cost	-0.070	-0.068 (-14.4) [0.5]	-0.067 (-14.2) [0.5]	-0.066 (-14.3) [0.9]	-0.066 (-14.3) [0.9]	-0.072 (-14.6) [-0.3]	-0.070 (-14.2) [0.0]	-0.071 (-14.5) [-0.1]	-0.070 (-14.4) [-0.0]	-0.069 (-14.4) [0.3]	-0.067 (-14.0) [0.6]	-0.070 (-14.5) [0.1]	-0.069 (-14.1) [0.3]	-0.074 (-14.5) [0.8]	-0.071 (-14.1) [-0.3]	-0.072 (-14.5) [-0.4]	-0.070 (-14.0) [0.1]
Travel Time	-0.150	-0.146 (-16.9) [0.4]	-0.145 (-16.1) [0.6]	-0.147 (-16.9) [0.4]	-0.147 (-16.8) [0.4]	-0.152 (-16.7) [-0.2]	-0.149 (-15.8) [0.2]	-0.151 (-16.7) [-0.1]	-0.148 (-16.7) [0.2]	-0.154 (-16.9) [-0.4]	-0.151 (-15.9) [-0.1]	-0.148 (-16.7) [0.2]	-0.145 (-15.7) [0.6]	-0.160 (-16.7) [-1.0]	-0.152 (-15.7) [-0.2]	-0.154 (-16.7) [-0.5]	-0.148 (-15.7) [0.2]
Access Time	-0.200	-0.196 (-11.2) [0.2]	-0.196 (-11.0) [0.2]	-0.196 (-11.3) [0.2]	-0.196 (-11.3) [0.2]	-0.202 (-11.1) [-0.1]	-0.199 (-11.0) [0.1]	-0.197 (-11.0) [0.2]	-0.202 (-11.1) [-0.1]	-0.205 (-11.4) [-0.3]	-0.202 (-11.4) [-0.1]	-0.194 (-11.0) [0.4]	-0.198 (-10.6) [0.1]	-0.209 (-11.2) [-0.5]	-0.211 (-11.1) [-0.6]	-0.198 (-10.9) [0.1]	-0.200 (-10.8) [0.0]
Mean Threshold *Cost*	0.08	-	0.001 (0.0) [0.0]	-	0.001 (0.0) [-3.5]	-	-	-	-	-	0.000 (0.0) [-0.8]	-	0.000 (0.0) [0.5]	-	-	-	0,004 (0.2) [-3.3]
Mean Threshold *Travel Time*	0.12	-	0.000 (0.3) [0.3]	-	-	-	0.110 (1.3) [0.1]	-	-	-	0.083 (5.4) [-2.4]	-	-	-	0.138 (1.3) [0.2]	-	0,096 (1.3) [-0.3]
Mean Threshold *Access Time*	0.16	-	0.018 (0.9) [0.9]	-	-	-		-	0.152 (2.1) [-0.1]	-	-	-	0.137 (1.1) [-0.2]	-	0.117 (1.6) [-0.6]	-	0,116 (0.7) [-0.3]
Log-likelihood		-546.4	-546.1	-552.8	-552.8	-519.3	-517.9	-522.9	-520.6	--529.2	-528.4	-531.6	-530,2	-498.1	-493.9	-510.1	-508,6
LR		0.5		0.0		2.6		4.5		1.6		2.7		8.5		6.6	
Critical χ^2 at 5% level		7.81		3.84		3.84		3.84		5.99		5.99		5.99		7.81	

Table A4. Models with simulated database, Case C, strong changes; n = 1,000

Threshold		Scenario															
		1		2		3		4		5		6		7		8	
Cost				√						√		√				√	
Travel Time						√				√				√		√	
Access Time								√				√		√		√	
Individuals changing their choice when threshold present		0		17		33		52		18		44		86		71	
Model	Target	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed	MNL	Proposed
Cost	-0.070	-0.071 (-17.0) [-0.1]	-0.071 (-15.6) [-0.1]	-0.071 (-17.1) [-0.1]	-0.072 (-16.8) [-0.4]	-0.073 (-16.9) [-0.6]	-0.072 (-16.5) [-0.4]	-0.073 (-16.4) [-0.6]	-0.071 (-16.3) [-0.1]	-0.071 (-16.9) [-0.3]	-0.071 (-16.7) [-0.1]	-0.071 (-16.4) [-0.1]	-0.069 (-16.1) [0.2]	-0.074 (-16.2) [-0.8]	-0.071 (-15.4) [-0.2]	-0.073 (-16.3) [-0.8]	-0.071 (-15.0) [-0.2]
Travel Time	-0.150	-0.138 (-15.2) [1.3]	-0.138 (-13.1) [1.1]	-0.138 (-15.2) [1.4]	-0.140 (-15.2) [1.1]	-0.143 (-15.3) [0.7]	-0.143 (-15.3) [0.8]	-0.141 (-15.1) [1.0]	-0.136 (-14.9) [1.6]	-0.141 (-15.3) [0.9]	-0.141 (-15.0) [1.0]	-0.138 (-15.0) [1.3]	-0.134 (-14.6) [1.7]	-0.148 (-15.2) [0.2]	-0.145 (-15.0) [0.5]	-0.146 (-15.2) [0.4]	-0.143 (-14.8) [0.7]
Access Time	-0.200	-0.206 (-15.0) [-0.4]	-0.206 (-14.1) [-0.4]	-0.211 (-15.4) [-0.8]	-0.210 (-15.2) [-0.7]	-0.201 (-14.6) [-0.1]	-0.204 (-14.5) [-0.3]	-0.174 (-13.1) [1.9]	-0.197 (-12.4) [0.2]	-0.202 (-14.8) [-0.1]	-0.203 (-14.7) [-0.2]	-0.171 (-13.0) [2.2]	-0.182 (-12.8) [1.3]	-0.170 (-12.7) [2.2]	-0.219 (-12.0) [-1.1]	-0.172 (-12.8) [2.1]	-0.213 (-12.7) [-0.7]
Mean Threshold *Cost*	0.08	-	0.000 (0.0) [0.0]	-	0.100 (1.4) [0.3]	-	-	-	-	-	0.000 (0.0) [-1.2]	-	0.017 (0.3) [-1.0]	-	-	-	0,036 (1.3) [-1.6]
Mean Threshold *Travel Time*	0.12	-	0.000 (0.0) [0.0]	-	-	-	0.079 (1.4) [-0.7]	-	-	-	0.032 (1.6) [-4.5]	-	-	-	0.061 (1.0) [-1.0]	-	0,040 (0.4) [-0.9]
Mean Threshold *Access Time*	0.16	-	0.000 (0.0) [0.0]	-	-	-	-	-	0.176 (5.0) [0.5]	-	-	-	0.107 (5.2) [-2.6]	-	0.332 (3.1) [1.6]	-	0,282 (9.4) [4.1]
Log-likelihood		-636.3	-636.3	-638.3	-637.3	-618.7	-617.4	-615.2	-612.5	-628.5	-628.4	-626.7	-625.5	-596.8	-585.5	-601.6	-595.5
LR		0.0		2.1		2.7		5.5		0.2		2.4		22.6		12.1	
Critical χ^2 at 5% level		7.81		3.84		3.84		3.84		5.99		5.99		5.99		7.81	