MODELLING INDIFFERENCE OPTION IN STATED PREFERENCE SURVEYS

JOHANNA AMAYA LEAL

Graduate Project as a requisite for graduation

Director

Ing. VICTOR CANTILLO MAZA, Ph.D

Universidad del Norte
Acreditada institucionalmente por su excelencia académica. Resolución MEN No 2085 - 05/09/2003.

School of Engineering
Industrial Engineering Department
Barranquilla, Colombia,
2007

ACKNOWLEDGEMENTS

I Desire to express my gratefulness to:

Victor Manuel Cantillo Maza, excellent director and guide in the reach of this project.

Carlos Daniel Paternina Arboleda, excellent boss and partner from whom I received support for the beginning and culmination of this project.

Correctors, whose valuable contributions enriched the final document,

The friends, who I knew during my studies and made of this a great experience.

CONTENTS

1 INTRODUCTION *6*

1.1 Thresholds in Choice Process **6**

1.2 Objectives **8**

1.3 Contributions **8**

1.4 Thesis Structure **9**

2 DISCRETE CHOICE MODELS *10*

2.1 General Modeling Assumptions **10**

2.1.1 Decision-maker 10

2.1.2 Alternatives 11

2.1.3 Attributes: 12

2.1.4 Decision Rule 12

2.2 Random Utility Theory **13**

2.2.1 Location and scale parameters: 14

2.2.2 Alternative specific constants: 15

2.2.3 The deterministic term of the utility 16

2.2.4 The random part of the utility 17

3 MODEL INCLUDING INDIFFERENCE *18*

4 EMPIRICAL ANALYSIS *26*

4.1 Application to Synthetic Data **26**

4.1.1 Subjective Values 29

4.1.2 Policies 29

4.2 Application to a Stated Preference Survey **31**

4.2.1 Experimental Design 32

4.2.2 Analysis and Results 32

4.2.3 Subjective Values 34

5 CONCLUSIONS *35*

REFERENCES *36*

APPENDIX A *39*

APPENDIX B *40*

APPENDIX C *41*

APPENDIX D *42*

LIST OF TABLES

Table 1. Parameters used for Attribute Generation
Table 2. Individuals whose choice differ according to the threshold influence
Table 3. Models with synthetic database
Table 4. Estimation of parameter ratios using the MNL and the proposed threshold model
Table 5. Policy Changes: percentage change in attribute values.
Table 6a. Comparison of simulated and modeled forecasts Scenario 2 δ ~ U[0, 0.5]
Table 6b. Comparison of simulated and modeled forecasts Scenario 3 δ ~ U[0, 1.0]
Table 7. Estimated models.
Table 8. Estimation of Marginal Substitution Rates (Col$/min).

LIST OF FIGURES

Figure 1a. Special choice probability cases given $\delta_q = 0$ …………………..22
Figure 1b. Special choice probability cases given $\delta_q = 1$,…………………..22
Figure 1c. Special choice probability cases given $\delta_q = 2$ …………………..22
Figure 2a. Choices probabilities as a function of $\delta_q = 0$ …………………..23
Figure 2b. Choices probabilities as a function of $\delta_q = 1$ ..…………………..23
Figure 2c. Choices probabilities as a function of $\delta_q = 2$ … ………………..23

MODELLING INDIFFERENCE OPTION IN STATED PREFERENCE SURVEYS

Abstract

An efficient choice design is characterized by utility balance, that is, the alternatives defined within each choice set have more or less equal choice probabilities. In these cases, respondents should be as nearly indifferent as possible. We propose a discrete choice model that describes the behaviour of individuals with indifference thresholds; this makes them perceive similar utilities from a specific situation in a stated preference (SP) experiment. The proposed model considers two alternatives plus the indifference option, so that the individual is not forced to choose one of the alternatives if s/he finds them equally attractive. We postulate that if thresholds exist they could be stochastic, random, differ among the population 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. As an illustration, the model was applied to synthetic and to real data coming from a stated preference survey; results show that when indifference thresholds exist in the population, the use of models without considering them can lead to errors in model estimation as well in prediction.

Keywords: Discrete choice models; Thresholds; SP survey; Modelling indifference

1 INTRODUCTION

Individual choices are affected by complex factors and it is a challenge to incorporate these in order to improve the realism of the modelling work. In particular, when confronting new situations people might undergo a process of experimentation and learning, through trial and error (Kitamura, 1990). This might trigger special conditions that individuals have to undergo in order to obtain utilities from the various processes that they might face.

1.1 Thresholds in Choice Process

The presence of thresholds in the perception and consideration of attributes is part of the complexity inherent to the choice-making problem (Cantillo and Ortúzar, 2006). In particular, the choice process may be affected by unknown restrictions on the perception and evaluation of attributes that can vary within the population according to their inherent characteristics (Cantillo et al, 2006).

The existence of thresholds, in general, has been acknowledged for many years. For example, in economics Quandt (1956) argued that there is a *band of indifference* around the level of utility derived from a commodity; and Georgescu-Roegen (1958) introduced the notion of thresholds into the theory of consumer choice, suggesting that a choice between any two commodities would be considered only if their quantities exceed some "*necessary minimum*"; thus, small increases in the quantity of products may leave consumers as well off as before, not affecting their utilities. In such a case, the thresholds of perception would be what Coombs et al. (1970) called "*just noticeable differences*".

In a seminal work, Krishnan (1977) formulated a model including thresholds as "*minimum perceivable differences*" between the utilities of the alternatives under consideration. According to this formulation an individual would prefer one alternative to another if the utility of the first exceeded that of the latter by at least a positive constant δ, the minimum perceivable difference. If, however, the excess of utility was less than the threshold level δ, a state of indifference would result. Krishnan incorporated the parameter δ into the conventional binary logit function and tested the derived model using data on transport mode choices. Based on this work, Lioukas (1984) found that thresholds could also be present when decision makers are confronted with more than two alternatives, and in those cases thresholds could lead to systematic and predictable intransitivity or indifference.

On the other hand, the new wave in stated preference (SP) methods postulates efficient choice designs that are characterized by utility balance (Huber and Zwerina, 1996), avoiding dominant alternatives. This involves defining alternatives with closer choice probabilities within each choice set, so utilities may end up easily being within an indifference threshold. In this context, we will consider thresholds in discrete choice model situations where the consumer is confronted with only two alternatives but also has the possibility of choosing a third, indifference, option. This way, the individual is not constrained to choose any of the alternatives, but s/he can choose an option of the type "both are the same". We developed an experimental design applying utility balance (Sandor, 2001) under the notion that if two alternatives are truly competitive the design is appropriate, and the individual responses can be within a threshold given the small difference between the alternative utilities.

1.2 Objectives

Supported in psychological foundations on the decision making, in the present investigation one looked for to conceptualize the process of election being considered the presence of indifference thresholds, and to develop specifications of framed mathematical models of discreet election in the theory of the random utility that they allow its incorporation. In order to obtain this objective it was necessary:

- Make a bibliographical revision respect to the existence of thresholds and the analysis of their behaviour.
- Develop a theoretical frame with respect to the decision making in the presence of thresholds.
- Propose a methodology to obtain the necessary data that it allowed considering models of discreet election considering thresholds.
- Apply proposed models to choice cases between transport alternatives, using simulated and real data to evaluate the impact to consider thresholds in estimation process of models and in the phase of prediction in front of changes coming from the application of diverse policies of transport.

1.3 Contributions

The present investigation tries to balance a combination of revision, synthesis and critical analysis of the psychological and methodological aspects for the incorporation of thresholds in the modeling of the election process, by means of the formulation and estimation of original models. In this order of ideas, the main contributions of this thesis can be synthesized as it follows:

• Development of a conceptual frame, with a solid psychological base, microeconomic consistency and statistical robustness, for the modeling of the election between discreet alternatives within the framework of the theory of the random utility incorporating the presence of thresholds.

• Preliminary analysis of the potential incidence of indifference thresholds in the estimation of marginal rates of substitution (p.e. the subjective value of the time) and in evaluation of projects.

• Study of the associated problem of incorrect specification to the use of discrete choice models without considering thresholds when there are evidences of their existence.

1.4 Thesis Structure

This document is organized as it is indicated next:

• Chapter 2 develops characteristics of discrete choice models and random utility theory.

• Chapter 3 presents the formulation of the model, its constraints and assumptions, and proposes procedures for its estimation.

• Chapter 4 presents two empirical analyses: one using simulated data and the other using data collected as part of a stated preference (SP) survey. In both cases we carry out a comparison of our method with the traditional compensatory model.

• Chapter 5 presents our conclusions, considering the implications of our threshold model in estimating the benefits of transport investments, and discusses some opportunities for further research.

2 DISCRETE CHOICE MODELS

We provide here a brief overview of the general framework of discrete choice models. We refer the reader to Ben-Akiva and Lerman (1985) for the detailed developments.

2.1 General Modeling Assumptions

The framework for a discrete choice model can be presented by a set of general assumptions. We distinguish among assumptions about the:

decision-maker -- defining the decision-making entity and its characteristics; alternatives -- determining the options available to the decision-maker; attributes -- measuring the benefits and costs of an alternative to the decision-maker; and decision rule -- describing the process used by the decision-maker to choose an alternative.

2.1.1 Decision-maker

Discrete choice models are also referred to as disaggregate models, meaning that the decision-maker is assumed to be an individual. The "individual" decision-making entity depends on the particular application. For instance, we may consider that a group of persons (a household or an organization, for example) is the decision-maker. In doing so, we may ignore all internal interactions within the group, and consider only the decisions of the group as a whole. We refer to "decision-maker" and "individual" interchangeably throughout this chapter. To explain the heterogeneity of preferences among decision-makers, a disaggregate model must include their characteristics such as the socio-economic variables of age, gender, education and income.

2.1.2 **Alternatives**

Analyzing individual decision making requires not only knowledge of what has been chosen, but also of what has not been chosen. Therefore, assumptions must be made about available options, or alternatives, that an individual considers during a choice process. The set of considered alternatives is called the choice set.

A discrete choice set contains a finite number of alternatives that can be explicitly listed. The choice of a travel mode is a typical example of a choice from a discrete choice set. The identification of the list of alternatives is a complex process usually referred to as *choice set generation*. The most widely used method for choice set generation uses deterministic criteria of alternative availability. For example, the possession of a driver's license determines the availability of the auto drive option.

The universal choice set contains all potential alternatives in the application's context. The choice set is the subset of the universal choice set considered by, or available to, a particular individual. Alternatives in the universal choice set that are not available to the individual are therefore excluded from the choice set.

In addition to availability, the decision-maker's awareness of the alternative could also affect the choice set. The behavioral aspects of awareness introduce uncertainty in modeling the choice set generation process and motivate the use of probabilistic choice set generation models that predict the probability of each feasible choice set within the universal set. A discrete choice model with a probabilistic choice set generation model is described later in this chapter as a special case of the latent class choice model.

2.1.3 **Attributes:**

Each alternative in the choice set is characterized by a set of attributes. Note that some attributes may be generic to all alternatives, and some may be alternative-specific.

An attribute is not necessarily a directly measurable quantity. It can be any function of available data. For example, instead of considering travel time as an attribute of a transportation mode, the logarithm of the travel time may be used, or the effect of out-of-pocket cost may be represented by the ratio between the out-of-pocket cost and the income of the individual. Alternative definitions of attributes as functions of available data must usually be tested to identify the most appropriate.

2.1.4 **Decision Rule**

The decision rule is the process used by the decision-maker to evaluate the attributes of the alternatives in the choice set and determine a choice. Most models used for travel behavior applications are based on *utility theory*, which assumes that the decision-maker's preference for an alternative is captured by a value, called utility, and the decision-maker selects the alternative in the choice set with the highest utility.

This concept, employed by consumer theory of micro-economics, presents strong limitations for practical applications. The underlying assumptions of this approach are often violated in decision-making experiments. The complexity of human behavior suggests that the decision rule should include a probabilistic dimension.

Some models assume that the decision rule is intrinsically probabilistic, and even complete knowledge of the problem would not overcome the uncertainty. Others consider the individuals' decision rules as deterministic, and motivate the uncertainty from the limited capability of the analyst to observe and capture all the dimensions of the choice process, due to its complexity.

Random utility models, used intensively in econometrics and in travel behavior analysis, are based on deterministic decision rules, where utilities are represented by random variables.

2.2 **Random Utility Theory**

Random utility models assume, as does the economic consumer theory, that the decision-maker has a perfect discrimination capability. However, the analyst is assumed to have incomplete information and, therefore, uncertainty must be taken into account. Manski (1977) identifies four different sources of uncertainty: unobserved alternative attributes; unobserved individual characteristics (also called "unobserved taste variations"); measurement errors; and proxy, or instrumental, variables.

The utility is modeled as a random variable in order to reflect this uncertainty. More specifically, the utility that individual *n* associates with alternative *i* in the choice set C_n is given by

$$U_{in} = V_{in} + \varepsilon_{in},$$

where V_{in} is the deterministic (or systematic) part of the utility, and ε_{in} is the random term, capturing the uncertainty. The alternative with the highest utility is chosen. Therefore, the probability that alternative *i* is chosen by decision-maker *n* from choice set C_n is

$$P(i|\, C_n) = P[\, U_{in} \geq U_{jn}\ \forall\, j \in C_n] = P[U_{in} = \max_{j \in C_n} U_{jn}].$$

In the following we introduce the assumptions necessary to make a random utility model operational.

2.2.1 **Location and scale parameters:**

Considering two arbitrary real numbers α and μ, where $\mu > 0$, we have that

$$P[\, U_{in} \geq U_{jn}\ \ \forall\, j \in C_n] =$$

$$P[\mu U_{in}+\alpha \geq\ \mu U_{jn}+\alpha\ \ \forall\, j \in C_n] =$$

$$P[\, U_{in} - U_{jn} \geq 0\ \forall\, j \in C_n].$$

The above illustrates the fact that only the signs of the *differences* between utilities are relevant here, and not utilities themselves. The concept of ordinal utility is relative and not absolute. In order to estimate and use a specific model arbitrary values have to be selected for α and μ. The selection of the scale parameter μ is usually based on a convenient normalization of one of the variances of the random terms. The location parameter α is usually set to zero. See also the discussion below of Alternative Specific Constants.

2.2.2 **Alternative specific constants:**

The means of the random terms can be assumed to be equal to any convenient value *c* (usually zero, or the Euler constant γ for Logit models). This is not a restrictive assumption. If we denote the mean of the error term of alternative *i* by $m_i = E[\varepsilon_{in}]$, we can define a new random variable $e_{in} = \varepsilon_{in} - m_i + c$ such that $E[e_{in}]=c$. We have

$$P[\, U_{in} \geq U_{jn} \;\forall\, j \in C_n] = P[\, V_{in} + m_i + e_{in} \geq V_{jn} + m_j + e_{jn} \;\forall\, j \in C_n],$$

a model in which the deterministic part of the utilities are $V_{in} + m_i$ and the random terms are e_{in} (with mean *c*). The terms m_i are then included as Alternative Specific Constants (ASC) that capture the means of the random terms. Therefore, we may assume without loss of generality that the error terms of random utility models have a constant mean *c* by including alternative specific constants in the deterministic part of the utility functions.

As only differences between utilities are relevant, only differences between ASCs are relevant as well. It is common practice to define the location parameter α as the negative of one of the ASCs. This is equivalent to constraining that ASC equal zero. From a modeling viewpoint, the choice of the particular alternative whose ASC is constrained is arbitrary. However, Bierlaire, Lotan and Toint (1997) have shown that the estimation process may be affected by this choice. In the context of the Multinomial Logit Model, they show that constraining the sum of ASCs to 1 is optimal for the speed of convergence of the estimation process. This result is also generalized for the Nested Logit Model.

2.2.3 **The deterministic term of the utility**

The deterministic term V_{in} of each alternative is a function of the attributes of the alternative itself and the characteristics of the decision-maker. That is

$$V_{in} = V(z_{in}, S_n)$$

where z_{in} is the vector of attributes as perceived by individual n for alternative i, and S_n is the vector of characteristics of individual n.

This formulation is simplified using any appropriate vector valued function h that defines a new vector of attributes from both z_{in} and S_n, that is:

$$x_{in} = h(z_{in}, S_n).$$

The choice of h is very general, and several forms may be tested to identify the best representation in a specific application. It is usually assumed to be continuous and monotonic in z_{in}. For a linear in the parameters utility specification, h must be a fully determined function (meaning that is does not contain unknown parameters). Then we have

$$V_{in} = V(x_{in}).$$

A linear in the parameters function is denoted as follows

$$V_{in} = \sum_k \beta_k x_{ink} \quad .$$

The deterministic term of the utility is therefore fully specified by the vector of parameters β.

2.2.4 **The random part of the utility**

Among the many potential models that can be derived for the random parts of the utility functions, we describe below the most popular. The models within the Logit family are based on a probability distribution function of the maximum of a series of random variables, introduced by Gumbel (1958). Probit and Probit-like models are based on the Normal distribution motivated by the Central Limit Theorem.

The main advantage of the Probit model is its ability to capture all correlations among alternatives. However, due to the high complexity of its formulation, very few applications have been developed. The Logit model has been much more popular, because of its tractability, but it imposes restrictions on the covariance structure. They may be unrealistic in some contexts. The derivation of other models in the "Logit family" is aimed at relaxing restrictions, while maintaining tractability.

We discuss here the specification and properties of the models from the Logit family (the Multinomial Logit model, the Nested Logit model, the Cross-Nested model and the Generalized Extreme Value model). After presenting the Probit model, we introduce more advanced models. The Generalized Factor Analytical Representation and the Hybrid Logit models are designed to bridge the gap between Logit and Probit models. The Latent Class Choice model is a further extension designed to explicitly include in the model discrete unobserved factors.

3 MODEL INCLUDING INDIFFERENCE

In the traditional binary random utility model, the probability that individual q chooses option A_i over A_j is given by (Ortúzar and Willumsen, 2001):

$$P(A_{iq}) = P\,(U_{iq} \geq U_{jq}) = P(V_{iq} + \varepsilon_{iq} \geq V_{jq} + \varepsilon_{jq}) \tag{1}$$

In this case the choice set is closed and the individual should choose either A_i or A_j.

Our proposed model considers the possibility of including a third, indifference, alternative characterised as "both are the same". Let us assume the existence of a non-negative indifference threshold δ_q, whereby if the absolute value of the utility difference is lower than this value, both alternatives will be undistinguishable for the individual; in that case, clearly the choice would be the indifference alternative. To work with δ_q it is necessary to make a few assumptions:

- δ_q is non-negative
- δ_q follows a probability distribution with density function $\phi(\delta)$ in the population.

With these, if A_i and A_j are the alternatives and U_i and U_j their respective utilities (assumed continuous functions over the possible values of their attributes), the following three choices are possible and individual q can choose among any of them:

(i) A_i is preferred to A_j if $U_{iq} > U_{jq} + \delta_q$

(ii) A_j is preferred to A_i if $U_{jq} > U_{iq} + \delta_q$

(iii) A_i or A_j are equally preferred if $\left| U_{iq} - U_{jq} \right| \leq \delta_q$

Under the new model the conditional probabilities to choose an alternative for a given δ_q are:

$$P(A_{iq} / \delta_q) = P(U_{iq} - U_{jq} \geq \delta_q) = P(U_{iq} \geq U_{jq} + \delta_q) \tag{2}$$

$$P(A_{jq} / \delta_q) = P(U_{jq} - U_{iq} \geq \delta_q) = P(U_{jq} \geq U_{iq} + \delta_q) \tag{3}$$

$$P(I_q / \delta_q) = P(\left| U_{jq} - U_{iq} \right| < \delta_q) \tag{4}$$

where $P(I_q / \delta_q)$ is the probability to choose "both are the same" given δ_q.

Thus, according with random utility theory we get:

$$P(A_{iq} / \delta_q) = P(\varepsilon_{jq} \leq \varepsilon_{iq} + V_{iq} - V_{jq} - \delta_q) \tag{5}$$

$$P(A_{iq} / \delta_q) = P(\varepsilon_{jq} \leq \varepsilon_{iq} + V_{iq} - (V_{jq} + \delta_q))$$

and

$$P(A_{jq} / \delta_q) = P(\varepsilon_{iq} \leq \varepsilon_{jq} + V_{jq} - (V_{iq} + \delta_q)) \tag{6}$$

$$P(I_q / \delta_q) = 1 - P(A_{iq} / \delta_q) - P(A_{jq} / \delta_q) \tag{7}$$

Solving Equation (7) to find the probability to choose the indifference alternative given a threshold, we get:

$$P(I_q / \delta_q) = P(V_{jq} - V_{iq} - \delta_q < \varepsilon_{iq} - \varepsilon_{jq} < V_{jq} - V_{iq} + \delta_q) \tag{8}$$

If we assume that the random errors ε are iid Gumbel distributed we get an explicit expression for the conditional probabilities given by the traditional logit formula:

$$P(A_{iq} / \delta_q) = \frac{e^{V_{iq}}}{e^{V_{iq}} + e^{V_{jq}+\delta_q}} \tag{9}$$

$$P(A_{jq} / \delta_q) = \frac{e^{jq}}{e^{V_{iq}+\delta_q} + e^{V_{jq}}} \tag{10}$$

$$P(I_q / \delta_q) = 1 - P(A_{iq} / \delta_q) - P(A_{jq} / \delta_q) \tag{11}$$

Note that the value of the threshold can be expressed as:

$$\delta_q = \overline{\delta} + \eta_q \text{,}$$

where $\overline{\delta}$ is an expected perceived value, common to all individuals of similar characteristics, and η_q represents an individual deviation following a certain density distribution function with mean zero and variance $\sigma^2{}_\eta$ [1].

If the population is homogenous, the distribution of η_q will be identical for all individuals *q*; nevertheless, it is possible to consider variations between individuals by expressing the mean values as a function of certain individual attributes (i.e. age, gender, income). For example, a linear-in-the-parameters expression might be used:

$$\overline{\delta_q} = \tau + \rho \mathrm{Y}_q$$

[1] A particular case is when $\sigma^2{}_n$=0, meaning that the threshold is a constant value in the population.

where τ and ρ are parameters and Y is a vector of socio-economic characteristics and choice conditions (i.e. trip purpose, travel frequency, possession of a travel card) among others.

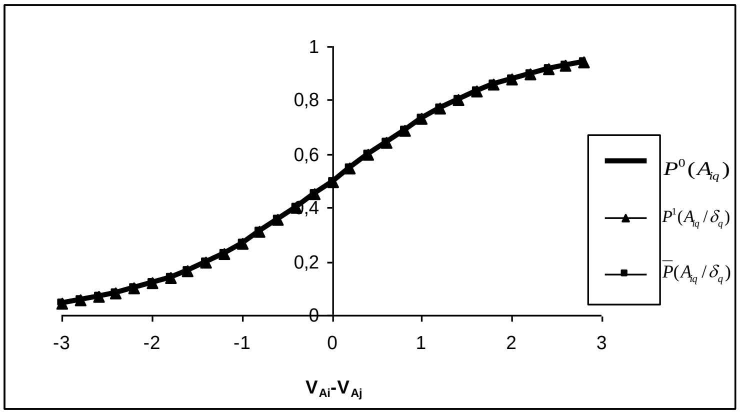
Figure 1 shows the variation of $P(A_{iq})$ for several values of δ_q. $P^0(A_{iq})$ describes the choice probabilities in the traditional binary logit model without indifference thresholds. The $P^1(A_{iq}/\delta_q)$ curve describes the choice probabilities according to the proposed model using δ_q. Finally, $\overline{P}(A_{iq}/\delta_q)$ was built as an adjusted probability in the following way[2]:

$$\overline{P}(A_{iq}/\delta_q) = P(A_{iq}/\delta_q) + P(I_q/\delta_q)\cdot 0.5 \qquad (12)$$

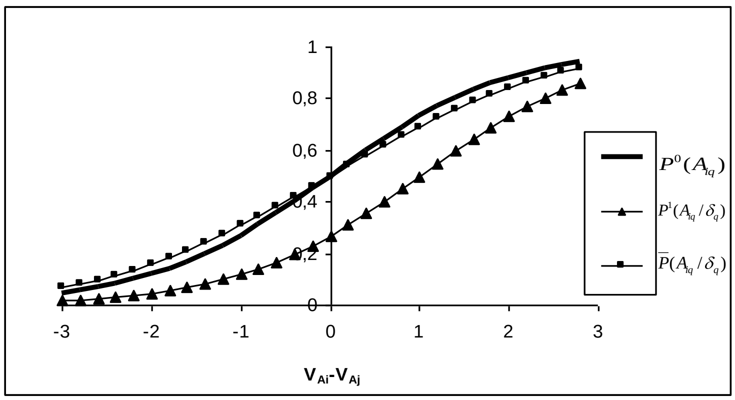
Obviously, the differences among the three curves increase with δ_q and, in fact, converge when δ_q is equal to zero.

Figure 2 illustrates the three choice probabilities (9), (10) and (11) for a given threshold δ_q. It can be noted that when the difference between $P(A_{iq}/\delta_q)$ and $P(A_{jq}/\delta_q)$ gets close to zero, $P(I_q/\delta_q)$ increases making apparent the presence of the threshold.

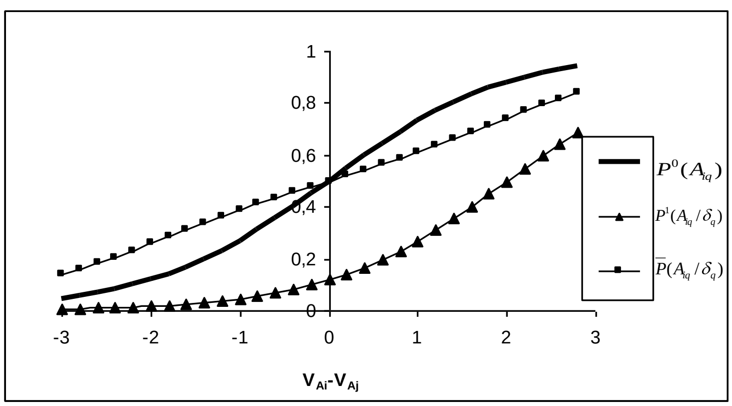
[2] This formulation could be used for forecasting as we avoid forcing individuals to choose when they are indifferent and try to give the same treatment to both alternatives when we include the new option.



a) $\delta_q = 0$

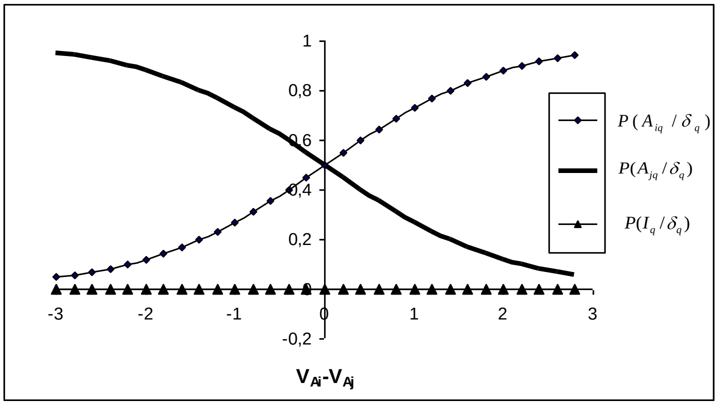


b) $\delta_q = 1$

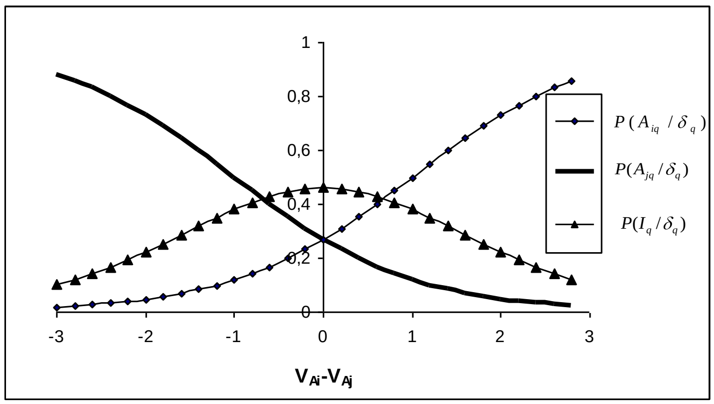


c) $\delta_q = 2$

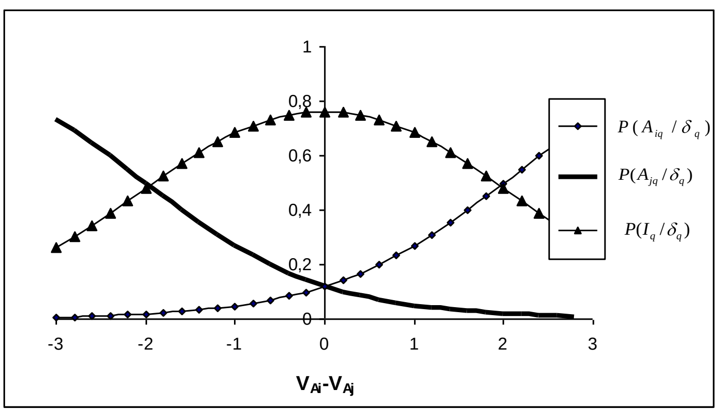
Figure 1: Special choice probability cases given δ_q



a) $\delta_q = 0$



b) $\delta_q = 1$



c) $\delta_q = 2$

Figure 2: Choices probabilities as a function of δ_q

To calculate unconditional probabilities it is necessary to integrate with respect to the distribution of δ as follows:

$$P(A_{iq}) = \int P(A_{iq} / \delta_q) \cdot \phi(\delta) d\delta$$

$$P(A_{jq}) = \int P(A_{jq} / \delta_q) \cdot \phi(\delta) d\delta \quad (13)$$

$$P(I_q) = \int P(I_q / \delta_q) \cdot \phi(\delta) d\delta$$

Equations (13) are difficult to solve analytically because they involve complex integrals with respect to the distribution $\phi(\delta)$[3]. However, we can use a probability simulator which leads to maximum simulated likelihood (MSL) estimates (Train, 2003). The estimator is determined in this case by the following equations:

$$\hat{P}(A_{iq}) = \frac{1}{R} \sum_{r=1}^{R} P(A_{iq} / \delta^r{}_q)$$

$$\hat{P}(A_{jq}) = \frac{1}{R} \sum_{r=1}^{R} P(A_{jq} / \delta^r{}_q) \quad (14)$$

$$\hat{P}(I_q) = \frac{1}{R} \sum_{r=1}^{R} P(I_q / \delta^r{}_q)$$

where $\delta^r{}_q$ denotes the r^{th} draws from the distribution of δ for individual q. To generate the set of discrete points we can use pseudo-random sequences (Silva and Garrido, 2003). So, incorporating the simulated probabilities, the log-likelihood is:

$$L(\delta) = \sum_{q=1}^{Q} \sum_{k} g_{qk} \log \hat{P}(A_{kq}), \quad (15)$$

[3] This mixed logit formulation is amenable to consider random parameters in the utility function.

where g_{qk} is equal to 1 if the individual chooses option k and 0 otherwise, k could be i, j or the indifference option.

It is important to emphasize that the traditional discrete choice model without thresholds (i.e. in this case a MNL model) is a particular case of our model when the means and variances of δ_q are zero (i.e. representing the absence of any thresholds).

4 EMPIRICAL ANALYSIS

We applied the proposed model to two data sets; the first is synthetic and the second coming from a SP survey including indifference option. Because the emphasis here was on thresholds rather than on the error term of the utility functions, in all cases we assume that errors ε are distributed iid Gumbel.

4.1 Application to Synthetic Data

To study the performance of the proposed model for a population where indifference thresholds exist, we worked in the same way as Williams and Ortúzar (1982), using the method of generating a simulated data. This consists of three hypothetical options: Bus, Shared Taxi and Indifference option, and include three attributes: Cost C, Travel time T_t and Access time T_a. Parameters used for attribute generation and coefficients in utility function are presented in Table 1.

Table 1: Parameters used for Attribute Generation

Attribute		**Initial Situation**		**Coefficient in Utility function (Target)**
		Bus	*Shared Taxi*	
Cost (C)	Mean	13	16	- 0.3
	Standard Deviation	1	1	
Travel Time (T_t)	Mean	35	20	- 0.040
	Standard Deviation	7	6	
Access Time (T_a)	Mean	13	10	- 0.080
	Standard Deviation	3	4	

Thresholds were also generated for each attribute following a uniform distribution, and we generated 4,000 independent observations according to the following procedure:

1. Initially we estimate the systematic utilities, error terms[4] and threshold according to the proposed model.

2. After procedure described in step 1, and according to utility function values, the individual chooses among the options, bus, shared taxi and the third one called "both are the same", following the compensatory utility maximizing process explained in section 2.

Table 2 shows the number of individuals who chose option 1 (bus), option 2 (shared taxi) or option 3 (indifference) according to the respective indifference threshold for each simulated case.

Table 2: Individuals whose choice differ according to the threshold influence

Option	***$\delta = 0$***	***$\delta \sim U[0,0.5]$***	***$\delta \sim U[0,1]$***	***$\delta \sim U[0,2]$***
Bus	1992	1779	1574	1222
Shared Taxi	2008	1783	1586	1199
Indifference		438	840	1579

We analyzed four scenarios, considering for each case the presence and absence of thresholds. For instance (see Table 3), first scenario consists of a population without thresholds influence. In next scenarios by contrast, are included thresholds values and we evaluate the influence introducing increases in their δ values. For each scenario, we estimated first a classical MNL model according to current conditions and no thresholds[5]; whereas the second one is the proposed model including thresholds with different values of δ following uniform probability function. Two *t-test* values appear: the first (in parenthesis) corresponds to the traditional null hypothesis $\theta_k = 0$, and the second one refers to the null hypothesis $\theta_k = \theta_v$, where θ_v is the true parameter value in the utility function of Table 1,

[4] In all cases we assumed that errors ε were distributed iid Gumbel with parameter scale = 1
[5] In case of indifference choices, we assigned them randomly to bus and shared taxi.

and is given as the target value in Table 3. On the other hand, b is the value of parameter threshold probability function considering δ ~ U [0,b].

Table 3: Models with synthetic database

		δ =0		δ ~ U[0, 0.5]		δ ~ U[0, 1.0]		δ ~ U[0, 2.0]	
Scenario		**1**		**2**		**3**		**4**	
Model	**Target**	**MNL**	**Proposed**	**MNL**	**Proposed**	**MNL**	**Proposed**	**MNL**	**Proposed**
Cost	-0.30	-0.273	-0.273	-0.266	-0.272	-0.248	-0.280	-0,235	-0.288
		(-15.7)	(-15.3)	(-15.4)	(-16.5)	(-14.5)	(-17.4)	(-13.9)	(-17.4)
		[-1.56]	[-0,02]	[-2.0]	[0,0]	[-3.0]	[0.0]	[-3.9]	[0.0]
Travel Time	-0.040	-0.040	-0.040	-0.038	-0.039	-0.038	-0.041	-0.033	-0.042
		(-12.8)	(-12.6)	(-12.1)	(-13.3)	(-12.3)	(-14.0)	(-10.9)	(-14.0)
		[-0.01]	[0.03]	[-0.8]	[-0.2]	[-0.6]	[0.2]	[-2.3]	[0.5]
Access Time	-0.080	-0.075	-0.075	-0.075	-0.076	-0.070	-0.080	-0,062	-0.080
		(-10.9)	(-10.8)	(-10.9)	(-11.5)	(-10.2)	(-12.5)	(-13.9)	(-12.2)
		[-0.69]	[-0.7]	[-0.7]	[-0.7]	[-1.5]	[0.0]	[-2.9]	[0.0]
b			0.000		0.482		0.949		1.937
			(0.229)		(21.79)		(30.50)		(39.72)
			[0.0]		[0,8]		[1,6]		[1,3]
$\bar{\rho}^2$		0.062	0.409	0.059	0.167	0.055	0.082	0.045	0.055
Log-Likelihood		-2597.03	-2597.15	-2605.88	-3660.39	-2617.51	-4033.28	-2644.68	-4151.36

The results for Scenario 1 show that in the absence of thresholds, the MNL and proposed models recovers parameters correctly, with estimates of threshold parameters equal to zero. In scenario 2, with a relatively small threshold value, the traditional MNL model it is not able to recover cost parameter, whilst proposed model recover it perfectly and have better fit[6]. In general, in scenarios with thresholds, parameter recovery is successful and the proposed model fits better than the traditional MNL model. In scenarios 3 and 4 it is evident the presence of mis-specification problem in traditional MNL model. It is obvious that at the same time threshold value increases, MNL models propend to mis-specification error while proposed model parameters are recovered accurately in all cases.

[6] Rho square adjusted is better fit index than log-likelihood keeping in mind that in MNL models there are two alternatives and in proposed model actually exists three alternatives to make an election; it could be noted that in all cases, these indices are better in proposed models.

4.1.1 Subjective Values

In order to investigate the potential influence of the indifference threshold on the model's subjective value of time point estimates, we also compared Travel time/cost and Access time/cost parameter ratios confirming that these were not affected by the scale parameter for either the MNL or the proposed threshold model (Table 4).

The 95% confidence intervals estimated by the proposed model for each of the parameter ratios include their respective target values in all cases. On the other hand, in the case of the MNL this is not true only for the Travel time/cost parameter ratio in scenario 3. The ratios Access time/cost parameters estimated by the MNL model include the target value in all scenarios.

Table 4: Estimation of parameter ratios using the MNL and the proposed threshold model

Parameter Ratio		*Travel time/cost*		*Access time/cost*	
Scenario	Target	13,33		26,67	
	Model	Point estimates	95% Confidence interval	Point estimates	95% Confidence interval
2	MNL	14,17	(12,61-15,74)	28,42	(23,19-33,65)
	Proposed	14,50	(13,04-15,96)	28,04	(23,17-32,91)
3	MNL	15,41	(13,71-17,10)	28,37	(22,78-33,95)
	Proposed	14,56	(13,16-15,95)	28,86	(24,21-33,51)
4	MNL	14,12	(12,36-15,88)	26,70	(22,25-31,15)
	Proposed	14,48	(13,09-15,87)	28,01	(23,39-32,63)

4.1.2 Policies

Following the approach of Munizaga et al. (2000) we tested four policies P1–P4; the specific changes in attributes associated with each one are shown in Table 5. Policies P1 and P2 correspond to small changes whilst P3 and P4 represent aggressive policy changes. The error measure considered was the percentage difference between the behaviour that was simulated for the modified attribute values estimated using the MNL model and that estimated using the threshold model.

Table 5: Policy Changes: percentage change in attribute values

Policy	*Cost Bus*	*Cost Shared Taxi*	*Travel Time Bus*	*Travel Time Shared Taxi*	*Access Time Bus*	*Access Time Shared Taxi*
P1			-10			
P2	+20					
P3			-20		-10	
P4	-10		-15		-15	

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

$$X^2 = \sum_i \frac{(\hat{N}_i - N_i)^2}{N_i} \qquad \textbf{(16)}$$

where $\hat{N}_i$ is the model estimate of the number of individuals choosing option $i(1 \le i \le 2)$ [7], and N_i is the first 1000 observations of simulated data bank. This result should be compared with the critical X^2 value at the 5% level with one degree of freedom, which is 3.84.

The results presented in Table 6 show that the threshold model proposed here always provides superior results in comparison to the MNL traditional model when the influence of threshold increases. When we apply the test to results in Scenario 2, δ ~ U [0, 0.5] (See Table 6a), in all cases, the X^2 test statistic for both models shows that the MNL errors were not significantly different from 0 at the 5% level. By contrast, if we apply the same to results in Scenario 3, δ~U[0,1.0] (See Table 6b), the MNL yields response errors larger than this value in two cases (P3 and P4). In all cases, estimations from proposed model are excellent while results from MNL model vary according to the magnitude of included changes in policies, in cases 1 and 2 this model performs well, and in cases 3 and 4 we note that specifications are out of expected level including significant errors in forecasting.

[7] We supposed individual's consumption is finally reflected in one of both: bus or shared taxi.

These results show that if there are relatively large indifference thresholds, the use of a model without thresholds can lead to prediction errors.

Table 6: Comparison of simulated and modeled forecasts

Policy	***Target***		***MNL***			***Proposed Model***		
	Bus	Shared Taxi	Bus	Shared Taxi	X^2	Bus	Shared Taxi	X^2
P1	543	457	533	467	0,20	532	468	0,24
			-1,8%	2,2%		-2,0%	2,4%	
P2	344	656	350	650	0,06	348	652	0,03
			1,6%	-0,9%		1,2%	-0,6%	
P3	595	405	585	415	0,20	585	415	0,20
			-1,7%	2,5%		-1,7%	2,5%	
P4	669	332	657	343	0,26	657	343	0,26
			-1,7%	3,5%		-1,7%	3,5%	

6a: Scenario 2 δ ~ U[0, 0.5]

Policy	***Target***		***MNL***			***Proposed Model***		
	Bus	Shared Taxi	Bus	Shared Taxi	X^2	Bus	Shared Taxi	X^2
P1	532	468	489	511	3,70	529	471	0,02
			-8,1%	9,2%		-0,6%	0,6%	
P2	349	652	343	657	0,06	347	653	0,00
			-1,6%	0,8%		-0,4%	0,2%	
P3	593	407	541	459	5,41	581	419	0,29
			-8,8%	12,8%		-2.0%	2,9%	
P4	664	337	596	404	9,11	653	347	0,22
			-10,2%	20,1%		-1,6%	3,1%	

6b: Scenario 3: δ = U[0, 1.0]

4.2 Application to a Stated Preference Survey

Data for this part is a result from SP survey applied in Barranquilla, Colombia, to individuals who did trips with work or study purpose. They should choose among three possible alternatives: Bus, Shared Taxi or “both are the same”.

4.2.1 Experimental Design

We designed the SP survey with choices based on four attributes: *Total cost*[8], *Travel time*, *Access time* (i.e. walking plus waiting time) and *Comfort*[9]. In the implementation of the experiment, respondents were asked to consider a trip they had recently taken, and then if they knew the alternatives. Each individual had to choose in nine hypothetical scenarios (See Appendix A, B and C) according to a fractional factorial design with balanced utilities to yield competitive alternatives allowing reflecting the influence of thresholds.

The design includes two variables in two levels and two variables in three levels.

Because of we get 36 observations including interactions, and keeping in mind that the purpose of analyze is focused on principal effects, we adopt a fractional designed following Kocur (1982).

To cross information of attributes of each alternative, the same was designed including the orthogonally constraint to guarantee an efficient design that reflect reality and balance between alternatives. (See Appendix D)

4.2.2 Analysis and Results

We interviewed 105 individuals getting 945 observations. Preliminary analysis of the data set showed that 11% of the sample was affected by the threshold's influence (104 observations), as they selected "both are the same". On the other hand, 14 respondents (i.e. 13,3% of the sample) answered the game lexicographically; of these, nine were lexicographic in the *Travel time* variable and five in the *Total cost* variable.

Four models were estimated (see Table 7, where t-test values appear in parenthesis). MNL1 is a model excluding the "both are the same" option, so it is based on only 844

[8] Cost in hundreds of Colombian pesos, where 1 US$ = Col$ 2,000 at the time of the survey.
[9] Dummy variable for Bus: equal to 1 if the vehicle had air condition and 0 otherwise.

observations. On the other hand, at the end of the game the individuals that had chosen the indifference option in some scenario were asked to choose again, but this time having to select one of the two alternatives as in the traditional binary model; we used this data to estimate model MNL2. Table 7 also presents models MT1 and MT2 that are consistent with our proposed approach (i.e. including the presence of indifference). In MT1 the threshold was taken as a constant value in the population $\delta_q = b$. Meanwhile, MT2 considers that δ_q follows a uniform distribution U [redacted]$),b$.

Our results show several things. First, but not very important, as the Bus penalty is negative it means that, *ceteris paribus*, people prefer Shared Taxi to Bus. Another general result is that, as it is normally the case with SP data, all parameters have the correct sign and are significant; interestingly this also happens with δ both when it is taken as a constant and when it follows a (albeit simple) probability distribution. In this latter sense it is obvious that models MT1 and MT2 are almost indistinguishable for this data set. However, if we compare the binary logit and the new models it is obvious that the threshold models fit the data better than the traditional MNL model, showing how people choices are affected by their thresholds in this SP experiment.

Table 7: Estimated models

Parameter	**MNL 1**	**MNL 2**	**MT1**	**MT2**
Travel cost	-0.305 (-4.05)	-0.280 (-3.94)	-0.259 (-3.92)	-0.260 (-3.92)
Travel time	-0.040 (-2.31)	-0.037 (-2.27)	-0.039 (-2.59)	-0.040 (-2.59)
Access time	-0.087 (-3.98)	-0.096 (-4.73)	-0.083 (-4.25)	-0.083 (-4.25)
Comfort	0.704 (4.42)	0.248 (2.27)	0.575 (4.16)	0.578 (4.15)
Bus constant	-0.984 (6.24)	-0.714 (5.42)	-0.825 (-6.26)	-0.829 (-6.26)

b			0.238 (10.72)	0.479 (10.62)
No. of observations	844	945	945	945
$\bar{\rho}^2$	0.06	0.05	0.16	0.16
Log-Likelihood	-544.6	-619.0	-871.2	-871.1

4.2.3 Subjective Values

Finally, and in order to investigate the potential influence of the indifference threshold on the model's subjective value of time point estimates, we compared the ratios of the parameters of Travel time and Cost and Access time and Cost (Table 8). As can be seen Access time is correctly (in the sense of, as usual) penalised nearly twice as much as Travel time. More interestingly, the subjective value of time point estimates are identical for both MNL models and very similar to those obtained from the MT models.

Table 8: Estimation of Marginal Substitution Rates (Col$/min)

Subjective Values	**MNL1**	**MNL2**	**MT1**	**MT2**
Travel time	13	13	15	15
Access time	28	34	32	32

5 CONCLUSIONS

The above analysis opens an interesting field in the area of model experimental design allowing to include and indifference option in SP surveys both at the theoretical and practical levels. We proposed a discrete choice model that incorporates random thresholds including the indifference alternative in a traditional binary choice context. Our formulation allows estimating the parameters of a threshold probability distribution starting from information about choices. The model calibration process can be done using simulated maximum likelihood.

The model was applied to synthetic data and then also 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 the magnitude of this depends on the magnitude of the thresholds (i.e. mis-especification errors increase with δ).

There are several interesting aspects for future research. One is to evaluate the implications of introducing marginal changes on the time and cost attributes under the assumption that an indifference threshold is in effect; in this case, willingness-to-pay could be null. Another aspect relates to generalizing the model and its application to different contexts, designing SP surveys with the “both are the same” option. Finally, it is important to develop more analyses and model designs including different threshold probability distributions.

REFERENCES

Ben-Akiva, M. E. and Lerman, S. R. 1985. Discrete Choice Analysis: Theory and Application to Travel Demand, MIT Press, Cambridge, Ma.

Ben-Akiva, M. and Bierlaire, M. (1999). Discrete choice methods and their applications to short-term travel decisions, in R. Hall (ed.), Handbook of Transportation Science, International Series in Operations Research and Management Science, Vol. 23 Kluwer.

Bierlaire, M., Lotan, T. and Toint, Ph. L. (1997). On the overspecification of multinomial and nested logit models due to alternative specific constants, Transportation Science 31(4): 363–371.

Cantillo, V., Ortúzar, J. de D., 2005. A semi-compensatory discrete choice model with explicit attribute thresholds of perception. Transportation Research 39B, 641–657.

Cantillo, V., Ortúzar, J. de D., 2006. Implications of thresholds in discrete choice modelling. Transport Reviews 26, 667–691.

Cantillo, V., Heydecker, B.G., Ortúzar, J. de D., 2006. A discrete choice model incorporating thresholds for perception in attribute values. Transportation Research 40B, 807-825.

Coombs, C.H., Dawes, R.M., Tversky, A., 1970. Mathematical Psychology: An Elementary Introduction. Prentice-Hall, Englewood 422 Cliffs, NJ.

Georgescu-Roegen, N., 1958. Thresholds in choice and the theory of demand. Econometrica 26, 157–168.

Huber, J., Zwerina, K., 1996. The importance of utility balance in efficient choice set designs. Journal of Marketing Research 33, 307-317.

Kitamura, R., 1990. Panel analysis in transportation planning: an overview. Transportation Research 24A, 401–415.

Krishnan, K.S., 1977. Incorporating thresholds of indifference in probabilistic choice models. Management Science 23, 1224–1233.

Kocur, G., T. Adler, W. Hyman y B. Aunet (1982)."Guide to forecasting travel demand with direct utulity assessment". Report Nª UMTA-NH-11-1-82, Urban Mass Transportation Administration, US Department of Transportation, Washington, D.C.

Lioukas, S.K., 1984. Thresholds and transitivity in stochastic consumer choice: multinomial logit analysis. Management Science 30, 110-122.

Manski, C. (1977). The structure of random utility models, Theory and Decision 8: 229–254.

Ortúzar, J. de D., Willumsen, L.G., 2001. Modelling Transport. John Wiley & Sons, Chichester.

Quandt, R.E., 1956. A Probabilistic Theory of Consumer Behaviour. Quarterly Journal of Economics 70, 507-536.

Sándor, Z., 2001. Computation, Efficiency and Endogeneity in Discrete Choice Models. PhD Thesis, University of Groningen.

Silva, M., Garrido, R.A., 2003. Estimation performance of low discrepancy sequences in stated preference. 10th International Conference on Travel Behaviour Research, Lucerne, Switzerland.

Train K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge.

Williams, H.C.W.L., Ortúzar, J. de D., 1982. Behavioural theories of dispersion and the mis-specification of travel demand models. Transportation Research 16B, 167–219.

APPENDIX A

PD SURVEY APPLIED TO SAMPLE

UNIVERSIDAD DEL NORTE

Universidad del Norte

División de Ingenierías

Encuesta a Usuarios de Bus Urbano/Taxi Colectivo

Acreditada institucionalmente por su excelencia académica. Resolución MEN No 2085 - 05/09/2003.

Responsable: Johanna Amaya Leal

1. IDENTIFICACIÓN DE LA ENCUESTA

Lugar ______ **Fecha** __ / __ / __ **Folio** ______

Encuestador ______ **Digitador** ______ **Hora** __ : __

2. IDENTIFICACIÓN DEL USUARIO

Ocupación

- Estudiante ☐
- Empleado ☐
- Independiente ☐
- Ama de Casa ☐
- Sin Ocupación ☐
- Otro ☐ → ¿Cuál? ______

¿Consideraría el bus/taxi colectivo como alternativa para realizar este viaje?

- Sí ☐
- No ☐ → **detener encuesta**

¿Ha utilizado anteriormente el bus/taxi colectivo como modo alternativo para hacer viajes al sitio de estudio o trabajo?

- Sí ☐
- No ☐ → **detener encuesta**

Sexo

- Femenino ☐
- Masculino ☐

Edad

- 15 - 20 AÑOS ☐
- 21 - 30 AÑOS ☐
- 31 - 40 AÑOS ☐
- 41 - 50 AÑOS ☐
- > 50 AÑOS ☐

3. INFORMACIÓN DEL VIAJE

¿Cuál es el motivo de su viaje?

- Trabajo ☐
- Estudio ☐
- Otro ☐

¿Con qué frecuencia realiza este viaje?

______ vez al(a) ______

¿Por qué escogió el bus/taxi colectivo para realizar su viaje?

______

¿Qué tipo de Bus utilizó?

- ☐ Con Aire Acondicionado
- ☐ Sin Aire Acondicionado

¿Por qué escogió este tipo de bus?

______

Hora de inicio del viaje (salida del origen)

__ : __

¿Cuál es el tiempo de duración de su viaje? *(estimado)*

horas ______ minutos ______

N/S ☐

¿Cuál fue el costo de su pasaje?

$ ______ N/S ☐

¿Cuál es su lugar de procedencia?

______

Hora de llegada de su viaje (arribo al lugar de destino)

__ : __

¿Cuál fue la duración de su caminata al paradero?

minutos ______

¿Cuál es su lugar de destino?

______

¿Cuánto tiempo esperó para realizar su viaje?

minutos ______

Si hubiera realizado este viaje en bus/taxi colectivo, ¿Cuál sería la duración de su viaje desde el lugar de origen hasta su destino final?

horas ______ minutos ______

N/S ☐

Si hubiera utilizado el modo alternativo, ¿Cuál sería el costo de su traslado desde el lugar de salida hasta su lugar de destino?

$ ______ N/S ☐

4. EVALUACIÓN DEL SERVICIO

Con nota de 1 a 10, ¿Cómo evaluaría los siguientes aspectos de estos modos de transporte?

	Bus	**Taxi colectivo**
La *frecuencia* del servicio	☐	☐
La *accesibilidad* al modo de transporte	☐	☐
La *comodidad* durante el viaje	☐	☐
La *seguridad* al viajar (en cuanto a accidentalidad)	☐	☐
La *seguridad* en el vehículo (en cuanto a posibilidad de robo o atraco)	☐	☐

5. PREFERENCIAS DECLARADAS

Considere el viaje que usted está por realizar. A continuación le mostraremos nueve situaciones en las que usted deberá escoger entre Bus y Taxi Colectivo. Ambos mantienen las mismas características que usted acaba de considerar en la evaluación del servicio (frecuencia,accesibilidad, comodidad, seguridad, etc.), pero cambian los tiempos de viaje y el costo del mismo.

PD N° ______

1	**Seguramente Bus** ☐	**Probablemente Bus** ☐	**Me da lo mismo** ☐	**Probablemente TC** ☐	**Seguramente TC** ☐
2	**Seguramente Bus** ☐	**Probablemente Bus** ☐	**Me da lo mismo** ☐	**Probablemente TC** ☐	**Seguramente TC** ☐
3	**Seguramente Bus** ☐	**Probablemente Bus** ☐	**Me da lo mismo** ☐	**Probablemente TC** ☐	**Seguramente TC** ☐
4	**Seguramente Bus** ☐	**Probablemente Bus** ☐	**Me da lo mismo** ☐	**Probablemente TC** ☐	**Seguramente TC** ☐
5	**Seguramente Bus** ☐	**Probablemente Bus** ☐	**Me da lo mismo** ☐	**Probablemente TC** ☐	**Seguramente TC** ☐
6	**Seguramente Bus** ☐	**Probablemente Bus** ☐	**Me da lo mismo** ☐	**Probablemente TC** ☐	**Seguramente TC** ☐
7	**Seguramente Bus** ☐	**Probablemente Bus** ☐	**Me da lo mismo** ☐	**Probablemente TC** ☐	**Seguramente TC** ☐
8	**Seguramente Bus** ☐	**Probablemente Bus** ☐	**Me da lo mismo** ☐	**Probablemente TC** ☐	**Seguramente TC** ☐
9	**Seguramente Bus** ☐	**Probablemente Bus** ☐	**Me da lo mismo** ☐	**Probablemente TC** ☐	**Seguramente TC** ☐

6. INGRESOS PERSONALES

¿Cuál es su ingreso personal? *(promedio mensual)*

Menos de $150.000 ☐	$300.001 - $450.000 ☐	$700.001 - $1.000.000 ☐	Más de $2.000.000 ☐
$150.001 - $300.000 ☐	$450.001 - $700.000 ☐	$1.000.001 - $2.000.000 ☐	No Contesta ☐

APPENDIX B

PD SURVEY APPLIED TO SAMPLE: SHORT TRIPS

Universidad del Norte

División de Ingenierías

Encuesta de Preferencias Declaradas

Usuarios de Buses Urbanos/Taxi Colectivo

Responsable: Johanna Amaya

PD1: **VIAJES CORTOS**

1

BUS SIN AA		TAXI COLECTIVO	
Costo	$ 1.200	Costo	$ 1,200
Tiempo	15 MIN	Tiempo	10 MIN
Acceso	10 MIN	Acceso	5 MIN

2

BUS SIN AA		TAXI COLECTIVO	
Costo	$ 1.200	Costo	$ 1.200
Tiempo	20 MIN	Tiempo	20 MIN
Acceso	10 MIN	Acceso	10 MIN

3

BUS CON AA		TAXI COLECTIVO	
Costo	$ 1.200	Costo	$ 1.200
Tiempo	25 MIN	Tiempo	30 MIN
Acceso	5 MIN	Acceso	10 MIN

4

BUS CON AA		TAXI COLECTIVO	
Costo	$ 1.200	Costo	$ 1.500
Tiempo	15 MIN	Tiempo	10 MIN
Acceso	10 MIN	Acceso	10 MIN

5

BUS SIN AA		TAXI COLECTIVO	
Costo	$ 1.300	Costo	$ 1.600
Tiempo	20 MIN	Tiempo	20 MIN
Acceso	5 MIN	Acceso	10 MIN

6

BUS SIN AA		TAXI COLECTIVO	
Costo	$ 1.300	Costo	$ 1,600
Tiempo	25 MIN	Tiempo	30 MIN
Acceso	10 MIN	Acceso	5 MIN

7

BUS SIN AA		TAXI COLECTIVO	
Costo	$ 1.300	Costo	$ 1.900
Tiempo	15 MIN	Tiempo	10 MIN
Acceso	5 MIN	Acceso	10 MIN

8

BUS CON AA		TAXI COLECTIVO	
Costo	$ 1.300	Costo	$ 1,900
Tiempo	20 MIN	Tiempo	20 MIN
Acceso	15 MIN	Acceso	10 MIN

9

BUS SIN AA		TAXI COLECTIVO	
Costo	$ 1.300	Costo	$ 1.900
Tiempo	25 MIN	Tiempo	30 MIN
Acceso	5 MIN	Acceso	5 MIN

APPENDIX C

PD SURVEY APPLIED TO SAMPLE: LONG TRIPS

UNIVERSIDAD DEL NORTE

Universidad del Norte

División de Ingenierías

Encuesta de Preferencias Declaradas

Usuarios de Buses Urbanos/Taxi Colectivo

Responsable: Johanna Amaya

PD2: **VIAJES LARGOS**

No.	Opción A		Opción B	
1	**BUS SIN AA**		**TAXI COLECTIVO**	
	Costo	$ 1.200	Costo	$ 1.200
	Tiempo	25 MIN	Tiempo	10 MIN
	Acceso	10 MIN	Acceso	5 MIN
2	**BUS SIN AA**		**TAXI COLECTIVO**	
	Costo	$ 1.200	Costo	$ 1.200
	Tiempo	30 MIN	Tiempo	25 MIN
	Acceso	10 MIN	Acceso	10 MIN
3	**BUS CON AA**		**TAXI COLECTIVO**	
	Costo	$ 1.200	Costo	$ 1.200
	Tiempo	35 MIN	Tiempo	40 MIN
	Acceso	5 MIN	Acceso	10 MIN
4	**BUS CON AA**		**TAXI COLECTIVO**	
	Costo	$ 1.200	Costo	$ 1.600
	Tiempo	25 MIN	Tiempo	10 MIN
	Acceso	10 MIN	Acceso	10 MIN
5	**BUS SIN AA**		**TAXI COLECTIVO**	
	Costo	$ 1.300	Costo	$ 1.700
	Tiempo	30 MIN	Tiempo	25 MIN
	Acceso	5 MIN	Acceso	10 MIN
6	**BUS SIN AA**		**TAXI COLECTIVO**	
	Costo	$ 1.300	Costo	$ 1.700
	Tiempo	35 MIN	Tiempo	40 MIN
	Acceso	10 MIN	Acceso	5 MIN
7	**BUS SIN AA**		**TAXI COLECTIVO**	
	Costo	$ 1.300	Costo	$ 2.100
	Tiempo	25 MIN	Tiempo	10 MIN
	Acceso	5 MIN	Acceso	10 MIN
8	**BUS CON AA**		**TAXI COLECTIVO**	
	Costo	$ 1.300	Costo	$ 2.100
	Tiempo	30 MIN	Tiempo	25 MIN
	Acceso	15 MIN	Acceso	10 MIN
9	**BUS SIN AA**		**TAXI COLECTIVO**	
	Costo	$ 1.300	Costo	$ 2.100
	Tiempo	35 MIN	Tiempo	40 MIN
	Acceso	5 MIN	Acceso	5 MIN

APPENDIX D

KOCUR VALUES AND ASSOCIATED ATTRIBUTES

	A	B	C	D	BUS A	BUS B	BUS C	BUS D
Set 1	0	0	0	0	1200	15	10	SIN AA
Set 2	0	1	1	0	1200	20	10	SIN AA
Set 3	0	2	2	1	1200	25	5	CON AA
Set 4	1	0	1	1	1200	15	10	CON AA
Set 5	1	1	2	0	1300	20	5	SIN AA
Set 6	1	2	0	0	1300	25	10	SIN AA
Set 7	2	0	2	0	1300	15	5	SIN AA
Set 8	2	1	0	1	1300	20	15	CON AA
Set 9	2	2	1	0	1300	25	5	SIN AA

	A	B	C	SHARED TAXI A	SHARED TAXI B	SHARED TAXI C
Set 1	0	0	0	1200	10	5
Set 2	0	1	1	1200	20	10
Set 3	0	2	2	1200	30	10
Set 4	1	0	1	1500	10	10
Set 5	1	1	2	1600	20	10
Set 6	1	2	0	1600	30	5
Set 7	2	0	2	1900	10	10
Set 8	2	1	0	1900	20	10
Set 9	2	2	1	1900	30	5