INFLUENCE OF DESIGN DIMENSIONS IN STATED CHOICE EXPERIMENTS: APPLICATION OF A HETEROSKEDASTIC LOGIT MODEL

Sebastián Caussade
Steer Davies Gleave Chile
Mariano Sánchez Fontecilla 310, piso 16, Santiago
Fax: 4736969; E-mail: Sebastian.Caussade@sdgworld.net

Juan de Dios Ortúzar
Department of Transport Engineering, Pontificia Universidad Católica de Chile
Casilla 306, Código 105, Santiago 22,Chile.
Fax: 5530281; E-mail: jos@ing.puc.cl

Luis Ignacio Rizzi
Department of Transport Engineering, Pontificia Universidad Católica de Chile
Casilla 306, Código 105, Santiago 22,Chile.
Fax: 5530281; E-mail: lrizzi@ing.puc.cl

ABSTRACT

We explore the influence of design complexity on the cognitive burden associated to stated choice experiments. Design complexity is analysed in terms of five dimensions, which were systematically varied according to a meta-experimental design. To study the influence of design complexity on respondent burden, we specified a heteroskedastic logit model with the scale parameter specified as a function of the design dimensions, namely: number of choice situations, number of available alternatives, number of attributes defining these alternatives, number of attribute levels and the range of the attribute levels. This allowed us to separate the effects of choice complexity from the marginal utility estimates. Our results show that all five design dimensions affect the choice variance, albeit in a very different manner; however, we did not find systematic effects on willingness-to-pay estimates.

1. INTRODUCTION

Stated choice (SC) methods include a variety of ways to elicit consumer's preferences in addition to the possibility of estimating willingness-to-pay (WTP) for improving specific attributes. For these reasons, the last decades have seen SC experiments become an accepted approach to study the behavioural response of consumers, households or even organizations in different applied fields such as transport, market research, health economics, etc (Louviere *et al*, 2000).

Still, even when choices provide useful information about consumer preferences, they contain what specialists have interpreted for years as noise or unexplained variance. Although modellers have developed methods to separate parameter estimates from noise, such efforts have tended to focus on the information supplied by the choices themselves ignoring the influence that the instrument design can have. This slant has arisen in spite of a growing body of research by both psychologists and specialists in decision making processes who have suggested that both task complexity and choice environment affect respondents facing complex situations (Swait and Adamowicz, 2001).

Given this evidence, the scope of this paper is to investigate, reveal and assess the influence (on model estimates and on measures of WTP) of the complexity of a stated route choice experiment carried out in Santiago de Chile. The complexity of the experiment was encompassed by the following design dimensions: number of available alternatives, number of attributes used to characterize the alternatives, number of choice situations presented to the respondent, number of attribute levels and variation range for those levels. These design dimensions are varied in a systematic fashion, according to a first experimental design hierarchy. As a second experimental design hierarchy, alternatives are characterised in terms of different travel time and travel cost attributes. The effect of complexity is modelled by means of a Heteroskedastic Logit (HL) model (Swait and Adamowicz, 2001; DeShazo and Fermo, 2002), that allows the scale parameter to be a function of the design dimensions. As to the structure of this paper, in Section 2 we do a literature review and postulate four hypotheses to be tested. In section 3 and section 4 we describe the experimental design and the modelling results. In section 5, we close the paper with the conclusions.

2. BACKGROUND LITERATURE AND HYPOTHESES

Simon (1955) questioned the rationality of human behaviour which was, until then, taken for granted. Further ideas were developed along these lines almost 30 years later by Heiner (1983), who incorporated the notion of information processing limitations in the consumer's ability to choose. In this context, Mazzotta and Opaluch (1995) empirically tested the validity of Heiner's hypothesis concerning choice complexity, and strongly suggested the existence of a gap between the cognitive ability of decision makers and the cognitive burden of the decision process.

Another source of literature comes from the behavioural decision theory field, including both theoretical and empirical research. The leading theories concerning the complexity of the decision environment are summarized by Payne *et al.* (1993). These and other authors have assessed how changes in the task environment impact the way people choose, leading to a wide spectrum of choice strategies, and suggest that the strategy selection depends on the trade-off between cognitive effort and outcome accuracy.

There are only a few authors who have treated this problem within a framework combining both suggestions above. Bradley and Daly (1994) were probably the first (certainly in transport) to model task complexity in a random utility framework. They used the logit scaling approach to test for fatigue effects in rank-order data concluding that the scale effect existed; this approach was also successfully used by Ortúzar and Rodríguez (2002) under fairly different circumstances. Saelensminde (2001) also implemented the scaling approach but to handle differences in the amount of unexplained variation due to inconsistencies in individual responses, showing that such a scaling effect also existed.

Swait and Adamowicz (2001) studied the problem in more depth by accounting for choice complexity and consumer behaviour through a parameterisation of the scale factor as a function of an entropy index associated to the experimental features. Simultaneously, DeShazo and Fermo (2002) examined both complexity and consistency through a parameterisation of the scale factor in terms of some measures that capture either the amount of information or the correlation structure of the data.

2.1 Testable Hypothesis

Several hypotheses on design complexity will be compared with the neoclassical null hypothesis that postulates that choice complexity does not have a significant effect on a respondent's ability to choose:

a) Number of alternatives per choice situation

Some studies (Wildert, 1998; DeShazo and Fermo, 2002; Arentze *et al.*, 2003) have proposed that having too many available options affects respondent's choice consistency. However, it may be argued that increasing the number of alternatives increases the probability that a respondent may find an option that matches her preferences better, leading to a more precise selection. Hence, our hypothesis is mixed in the sense that an initial increase on the number of options may decrease the variance of the error term up to a threshold; but if that number continues to grow, the variance will start to increase. This would suggest a U-shaped relationship between the number of options and the variance of the error term.

b) Number of attributes per alternative

Our hypothesis is that an increase in the number of attributes will always produce an increase in the variance of the error term. This may hold because as respondents attempt to process more information, they can either make mistakes or adopt a simplifying strategy (heuristic) based only on partial information (Arentze *et al,* 2003). Both aspects would compromise the consistency of the decision process by increasing the variance of the error term (Hensher, 2003).

c) Number of levels per attribute and range of attribute variations

Based on the results of Dellaert *et al.* (1999) and Ohler *et al.* (2000), we propose that experimental complexity should increase as the number of attribute levels grows, simply because a larger number of comparisons have to be made. In turn, the variance of the error term should be minimised in the presence of a narrow range of variation, because comparisons would be easier to assess, leading to a more consistent process.

d) Number of choice situations to be assessed

This is probably the most controversial design dimension. Stopher and Hensher (2000) discussed the influence of the number of choice situations to be assessed, concluding that it had a marginal effect on consistency. However, Bradley and Daly (1994) found that increasing the number of choice situations to be evaluated led to an increase in the error term variance because the fatigue effects built up; this effect was also found by Ortúzar *et al* (2000). Thus, we postulate a mixed effect in which an initial increase in the number of choice situations may decrease the variance of the process up to a threshold (because of a learning effect); after this, the error variance should increase as the number of choice situations continues to grow because of the cognitive burden overload.

3. EXPERIMENTAL DESIGN

The only way to assess the influence of experimental complexity is to allow for systematic variations in the design dimensions of a SC experiment. This requirement represents an interesting challenge because modellers usually define the design dimension values before starting the experimental design, and maintain them fixed across the entire design.

In this study we used the Design of Designs (DoD) SC software originally developed and applied in Sydney (Hensher, 2003) which allows the analyst to build different survey designs, each one presenting respondents some differences in terms of one or more design dimensions. The approach assumes in a first stage that the dimensions are themselves the attributes of the experiment. Thus, it is possible to construct different designs through the manipulation of those dimensions by experimental design principles. The values that each design dimension can take in our route choice experiment are presented in Table 1.

Table 1: Dimensionality of the design plan

Choice set size	Number of alternatives	Number of attributes	Number of attribute levels	Range of attribute levels
6	3	3	2	Narrower than base
9	4	4	3	Base
12	5	5	4	Wider than base
15	----	6	----	----

This way, 16 different designs can be built to test the impact of each of the five design dimensions. The features for each design are presented in Table 2. The 16 designs are computer-generated in order to reduce the correlations between attributes and to maximize the amount of information captured by each choice task. The design developed herein takes into account the expected signs of the parameters (e.g. negative for the time and cost attributes). Hence, the search eliminates dominant alternatives, which do not provide useful information for estimation.

For the second stage six route choice attributes were selected to characterize the options: free-flow time, slowed down time, stop/start time, variability of trip time, toll cost and running costs. The latter are based on the trip distance reported by the respondent when asked about a recent trip of certain conditions, and on average fuel consumption.

Table 2: The sub-designs of the overall design

Choice set of size	Number of alternatives	Number of attributes	Number of levels of attributes	Range of attribute levels
15	4	4	3	Base
12	4	4	4	Wider than base
15	3	5	2	Wider than base
9	3	5	4	Base
6	3	3	3	Wider than base
15	3	3	4	Narrower than base
6	4	6	2	Narrower than base
9	5	3	4	Wider than base
15	5	6	4	Base
6	5	6	3	Wider than base
6	4	5	4	Narrower than base
9	5	4	2	Narrower than base
12	4	6	2	Base
12	3	3	3	Narrower than base
9	3	4	2	Base
12	5	5	3	Narrower than base

To explore how varying the number of attributes affects subjective values, the attributes were grouped according to the following patterns:

- *designs concerning three attributes:* total time (free flow + slowed down + stop/start time), trip time variability, total costs (toll + running cost)
- *designs concerning four attributes:* free flow time, congestion time (slowed down + stop/start), trip time variability, total costs
- *designs concerning five attributes:* free flow time, slowed down time, stop/start time, trip time variability, total costs
- *designs concerning six attributes:* free flow time, slowed down time, stop/start time, trip time variability, toll cost, running cost.

In the implementation of the experiment, respondent were first asked about a trip they had recently done and had to report the trip attributes required by the DoD they were assigned to; we called this trip Current Route ("Ruta Reciente" in Spanish). The computer program automatically generated the hypothetical choice scenarios. Each specific design pivoted from the attribute levels associated with the Current Route. Figure 1 shows a typical SC screenshot.

Each hypothetical choice situation could yield two observations. First the individual had to choose among her current route ("Viaje Reciente") and the alternatives ("Ruta Alternativa" B, C, D, and even E in the most complex case) on the basis of the attributes described on the screen, leading thus to a second choice only among the alternative routes.

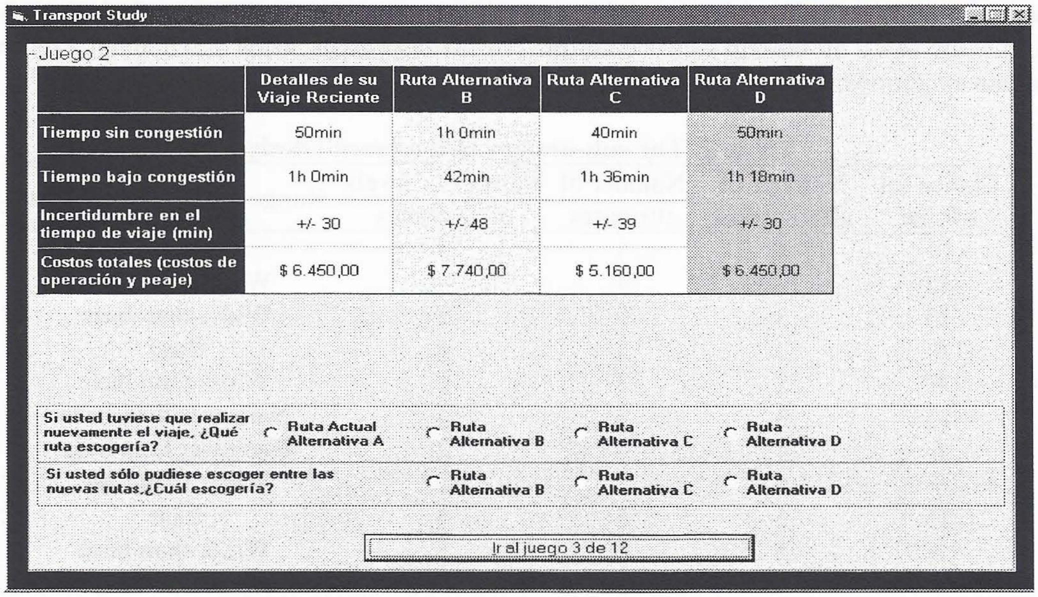

Figure 1: An example of a stated choice screen

4. MODELLING APPROACH

The logit family of models are based on random utility theory (Ortúzar and Willumsen, 2001); this basically postulates that individuals belong to a given homogeneous population Q, act rationally and possess perfect information, i.e. they always select that option which maximizes their net personal utility subject to legal, social, physical and/or budgetary constraints. It is also postulated that there is a certain set $A = \{A_1, A_2, \ldots, A_N\}$ of available alternatives and a set X of vectors of measured attributes of the individuals and their alternatives. Each $A_j \in A$ has associated a net utility U_{iq} for individual q. The modeller does not possess complete information concerning the individuals, therefore he assumes that U_{iq} can be represented by the summation of two components:

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

The first is a measurable or systematic part V_{iq} which is a function of the attributes, while the random part ξ_{iq} reflects the idiosyncrasies and particular tastes of each individual. The systematic component V_{iq} is generally assumed to be additive and linear in the fixed marginal utility parameters (2):

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

In turn, and without loss of generality, it can be assumed that the residuals ξ are random variables with mean zero and a certain probability distribution to be specified. If the random error terms are distributed IID Gumbel the specification yields a Multinomial Logit Model (MNL), which is the simplest and most popular discrete choice model and whose probability expression is given by (Domencich and McFadden, 1975):

$$P_{iq} = \frac{e^{\lambda \cdot V_{iq}}}{\sum_{A_j \in A(q)} e^{\lambda \cdot V_{jq}}} \tag{3}$$

where the utility functions are as in (2) and the parameter λ is a scale factor inversely proportional to the common standard deviation of the Gumbel error terms (4):

$$\lambda = \frac{\pi}{\sqrt{6} \cdot \sigma} \tag{4}$$

As it can be inferred from (4), given that the error terms are IID there is no possibility for the scale factor to vary across individuals, so it represents a constant term. Also, as it cannot be estimated separately form the vector of parameters $\bar{\theta}$, it is normalised in practical applications.

4.1 The Heteroskedastic Logit Framework

The heteroskedastic (or covariance heterogeneity) logit (HL) model allows the analyst to incorporate the complexity and cognitive burden of the experiment through an appropriate parameterisation of the scale factor. Noting that the scale factor is inversely proportional to the variance, we would expect that choice dimensions which lead to more complex decision choice processes would induce higher error rates and thus lower scale parameters. DeShazo and Fermo (2001) proposed an exponential functional form for λ to preclude negative scale parameters:

$$\lambda_k(C_k) = \exp\left[\sum_{l=1}^{5} \gamma_l \cdot C_l\right] \tag{5}$$

Here the scale factor is no longer a constant term. Instead, it is a function of the complexity of the experiment *(C_k)* which is in turn approximated by the five design dimensions; this can give raise to three types of heteroskedasticity. First, the scale parameter differs across responses from different DoD since systematic variation is present across design dimension variables. Second, the scale parameter differs across individual responses within each DoD, since it is also a function of the number of choice situations already considered. Third, socio-economic variables can be introduced in (5) to obtain a fully heteroskedastic model, i.e. the *n*th choice situation within a DoD for two individuals will have a scale parameter that varies according to their socio-economic variables. Incorporating the parameterisation of the scale factor (5) in the choice probability (3) leads to the following HL expression for the probability of choice:

$$P_{iqk} = \frac{e^{\lambda_k(C_k) \cdot V_{iqk}}}{\sum_{A_j \in A(q)} e^{\lambda_k(C_k) \cdot V_{jqk}}} \tag{6}$$

Once again, the scale parameter cannot be identified and has to be normalized assuming no constant term. If all the parameters γ_l in (5) turn out to be zero, then λ equals one and we obtain the MNL model which is a restricted version of (6). To estimate the HL model a code was written in GAUSS using the MAXLIK routine to maximize the values of the log-likelihood function.

5. MODEL ESTIMATIONS

The final survey was carried out in Santiago, Chile during July 2003. The sample consisted of 403 individuals who generated 8020 valid responses. Several HL models were specified in addition to their corresponding MNL. We only show results for our best specifications. Table 3 contains the estimates of the trip attribute variables that were already described in section 3 plus an added binary variable 'Current' (an inertia variable) that takes the value of one for the current route and zero otherwise.

All parameters are statistically significant and have the expected *a priori* sign in both cases. The only variable that is not statistically significant is Current, suggesting there is no inertia effect. The table also shows the log-likelihood of the MNL and HL models; a-likelihood ratio test (LR) strongly rejects the homoskedastic MNL form (i.e. LR equals 545.3 and the critical $\chi^2_{11}(0.99)$ is 24.7). However, note that the MNL is based here on the strong assumption that data from each different design can simply be joined together for estimation.

Table 3: Estimation results for the MNL and HL models

Coefficients	MNL Model	t-ratios	HL Model	t-ratios
Current (0, 1)	0.0066	0.17	0.0110	0.03
Free-flow time (min)	-0.0122	-22.56	-0.0254	-4.59
Slow down time (min)	-0.0154	-9.63	-0.0296	-4.01
Stop/start time (min)	-0.0425	-19.81	-0.0732	-4.22
Congested time (min)	-0.0340	-14.70	-0.0455	-4.69
Total time (min)	-0.0199	-24.20	-0.0174	-4.70
Travel time variability (min)	-0.0051	-10.32	-0.0088	-4.48
Running costs (Ch$)	-1.00E-05	-2.96	-1.00E-04	-2.55
Toll costs (Ch$)	-1.00E-04	-4.85	-4.00E-04	-3.53
Total costs (Ch$)	-2.00E-04	-22.85	-2.0E-04	-4.65
Log-likelihood (C)	-9294.98		-9294.98	
Log-likelihood (θ)	-8131.15		-7845.11	
ρ^2 (C)	0.125		0.156	

Based on 8020 observations

On the other hand, separate MNL for each design were also estimated and the sum of their log-likelihood values at convergence was higher than the log-likelihood of the HL model (i.e. -7599.66 against -7845.11 for an additional 70 parameters). As these two models are non-nested we can not compute a χ^2 test, but the result would cast doubts on the validity of our HL model, implying that MNL models reflect not only differences across designs but also taste differences. As we took special care in selecting a sample of individuals as homogenous as possible, we feel that taste variability mainly arises out of insufficient sample size for each design and hence we maintain the HL hypothesis. Table 4 shows estimates of the design dimension variables.

The number of attributes has clear detrimental effect on the ability to choose, contributing to a higher error variance. The same is true for the number of levels, although the negative effect on variance is approximately three times smaller.

Table 4: Coefficient estimates for the parameterisation of λ

Attribute	HL Model	t-ratios
Number of choice situations	0.056	2.22
Squared number of choice situations	-0.0033	-1.78
Number of attribute levels	-0.122	-2.71
Narrow range	0.894	10.61
Wide range	-0.284	-3.32
Dummy 4 alternatives	0.609	6.09
Dummy 5 alternatives	0.485	6.51
Dummy 4 attributes	-0.583	-4.91
Dummy 5 attributes	-1.012	-9.06
Dummy 6 attributes	-1.456	-10.34
Income	0.019	1.84

Based on 8020 observations

The effect of the range, however, differs depending on if it is narrow or wide. The former contributes to an important reduction in variance, whereas the wide range contributes to a higher variance. From this we are tempted to conclude that narrow ranges place less cognitive burden on respondents, since trade-offs among attributes tend to be similar across responses. With respect to the number of alternatives an inverted U-shaped pattern emerged. Designs with four alternatives possess the highest scale parameter and designs with five alternatives possess a lower scale parameter but higher than that corresponding to three alternatives. This evidence lends some support to our original hypothesis (a) in section 2.2. A quadratic relationship holds for the effect of the number of situations. SC experiments with nine or ten choice situations seem to be optimal in terms of minimising error variance, supporting the stated hypothesis. Although it is impossible to disentangle the learning effect from the cumulative burden effect in this relationship, learning effects tend to dominate up to ten different choices, whereas cognitive burden effects seem to prevail from ten choices onwards.

6. DERIVING VALUES FOR TRAVEL TIME SAVINGS (VTTS)

To estimate the willingness-to-pay (WTP) for saving travel time modellers need the marginal rate of substitution between travel time and cost. In the case of linear-in-parameters utility functions, the subjective value of time equals the ratio between the coefficients of travel time (t_i) and cost (C_i) at constant utility level (Gaudry *et al*, 1989):

$$VTTS = \left.\frac{\theta_t}{\theta_C}\right|_{U=cte} \tag{8}$$

The specification of our experiment allowed for disaggregating total travel time into subcomponents that considered different travel conditions (section 3.1). This enabled us to estimate different subjective values for the various travel time components (Table 5).

WTP for free-flow time, slow down time, stop/start time and congestion time were derived using the toll cost parameter as the denominator in (8). However, WTP for total travel time savings had to be derived using the total cost parameter because all sub-designs including total travel time also involved the aggregate variable total cost.

Table 5: Subjective values of time

VTTS [US$/min]	MNL Model	Confidence Interval	Heteroskedastic Model	Confidence Interval
Free flow time	0,19	[0,14 - 0,32]	0,20	[0,18 - 0,22]
Slow down time	0,24	[0,16 - 0,42]	0,23	[0,18 - 0,29]
Stop/start time	0,67	[0,48 - 1,13]	0,58	[0,48 - 0,68]
Congestion time	0,54	[0,37 - 0,91]	0,36	[0,32 - 0,41]
Total time	0,16	[0,14 - 0,18]	0,14	[0,12 - 0,16]

* At the time of the survey, 1 US$ = 630 Ch$.

The values for the different travel time components seem plausible. Free-flow time is the least onerous and stop/start time the most expensive one; congested time is also perceived as relatively expensive. These results tell us that to increase social welfare it is better to reduce one minute of congested travel time than one minute of free-flow travel time.

VTTS point estimates tend to be slightly higher for the pooled MNL model; however, for most travel time components, point estimates are included within the corresponding confidence interval for the same attribute in the alternative model, with the exception of congestion time [8]. This result is different to that found by Hensher (2001, 2004a), DeShazo and Fermo (2002) and Saelesminde (2001), but it is in good agreement with previous results in Chile (Gaudry *et al*, 1989) and elsewhere (Carlsson, 2003; Train, 1998). So there is no clear agreement on this issue.

7. CONCLUSIONS AND EXTENSIONS

The main purpose of this paper was to investigate the influence of SC design complexity on consumer's ability to choose. For this we conducted a stated route choice experiment proceeding in a non-conventional fashion. This way we were able to estimate a heteroskedastic logit model, where the Gumbel scale parameter (inversely related to its error variance) was parameterised in terms of the design variables and was allowed to vary across choice scenarios.

We found that the number of attributes had a clear detrimental effect on the ability to choose, contributing to a higher error variance. The number of levels had a much smaller negative effect. The effect of the range differed depending on if it was narrow or wide, the former contributing to an important reduction in variance. The number of alternatives gave an inverted U-shaped pattern with the optimum for four alternatives. Finally, although the number of choice situations also gave a U-shaped pattern (with an optimum around 9-10), its importance was smaller than the rest of the design dimensions.

We were able to separate total travel time into different components according to the level of congestion experienced on a route. A clear pattern emerged, the more congested the route is the higher is the WTP for saving travel time. This result could have implications for the design of

Advance Traveller Information Systems (ATIS) intended to provide drivers with on-line on-route information: drivers would be willing to pay more for avoiding extreme congestion than for reducing travel time in free-flow or low congestion conditions.

A first extension to our work would be to seek different parameterisations of the scale factor: information theory could provide us with some index of choice complexity aside from the one considered by Swait and Adamowicz (2001). We could also compare our HL model with a model estimated allowing for different scale factors associated to each design; that model should be at least superior to the simple MNL used as reference here. Another extension could be to consider a heteroskedastic-mixed logit specification, where attribute parameters are randomly distributed in the population and the Gumbel scale parameters are functions of the design dimensions.

REFERENCES

Bradley, M. y A.J. Daly (1994) Use of the logit scaling approach to test rank-order and fatigue effects in stated preference data. **Transportation 21**, 167-184.

Carlsson, F. y P. Martinsson (2003) Design techniques for stated preference methods. **Health Economics 12**, 281-294.

Caussade, S. (2003) Complejidad Experimental y Confiabilidad de los Resultados en Encuestas de Preferencias Declaradas. MSc Thesis, Department of Transport Engineering, Pontificia Universidad Católica de Chile.

Dellaert, B., J. Brazell y J.J. Louviere (1999) The effect of attribute variation on consumer choice consistency. **Marketing Letters 10**, 139-147.

De Shazo, J.R. y G. Fermo (2002) Designing choice sets for stated preference methods: the effects of complexity on choice consistency. **Journal of Environmental Economics and Management 44**, 123-143.

Domencich, T.A. y D. McFadden (1975) **Urban Travel Demand: A Behavioural Analysis**. North Holland, Amsterdam.

Gaudry, M.J.I., S.R. Jara-Díaz y J. de D. Ortúzar (1989) Value of time sensitivity to model specification. **Transportation Research 23B**, 151-158.

Heiner, R. (1983) The origin of predictable behaviour. **The American Economic Review 73,** 560-595.

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

Mazzotta, M. y J. Opaluch (1995) Decision making when choices are complex: a test of Heiner's hypothesis. **Land Economics 71**, 500-515.

Ohler, T., A. Le, J.J. Louviere y J.D. Swait (2000) Attribute range effects in binary response tasks. **Marketing Letters 11**, 249-260.

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

Payne, J., J. Bettman, y E. Johnson (1993) **The Adaptive Decision Maker**. Cambridge University Press, Cambridge.

Saelensminde, K. (2001) Inconsistent choices in stated choice data: use of the logit scaling approach to handle resulting variance increases. **Transportation 28,** 269-296.

Simon, H.A. (1955) A behavioural model of rational choice. **Quarterly Journal of Economics 69**, 99-118.

Stopher, P.R., D.A. Hensher (2000) Are more profiles better than fewer? Searching for parsimony and relevance in stated choice experiments. **Transportation Research Record 1719**, 165-174.

Swait, J.D., W. Adamowicz (2001) Choice environment, market complexity and consumer behaviour: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice. **Organizational Behaviour and Human Decision Processes 86**, 141-167.

Widlert, S. (1998) Stated preference studies: the design affects the results. In J. de D Ortúzar, D.A. Hensher and S.R. Jara-Díaz (eds.), **Travel Behaviour Research: Updating the State of Play**, Pergamon. Oxford.