WORKING PAPER
ITS-WP-04-14

How do Respondents Handle Stated Choice Experiments? – Information processing strategies under varying information load

By

David A. Hensher

May, 2004

ISSN 1440-3501

INSTITUTE OF TRANSPORT STUDIES

The Australian Key Centre in Transport Management

The University of Sydney and Monash University

Established under the Australian Research Council's Key Centre Program.

NUMBER:	Working Paper ITS-WP-04-14
TITLE:	How do Respondents Handle Stated Choice Experiments? – Information Processing Strategies under Varying Information Load

ABSTRACT:

The popularity of stated choice (SC) experiments has spurned a large number of design strategies within which to study choice behaviour. When the amount of information provided increases, we often wonder how an individual handles such information in making a choice. Defining the amount of information (or 'complexity') as the product of the number of attributes and number of alternatives associated with each choice set, we investigate how this information is processed as we vary the amount of information. Four ordered heterogeneous logit and mixed logit models are developed, each for a fixed–attribute design, in which the dependent variable is the difference between the maximum (fixed) number of attributes in the design and the actual number that were maintained by the respondent in their information processing strategy (IPS). We have found that individuals adopt a range of 'coping' or editing strategies that are consistent with how we normally process information in real markets. Importantly, we should not argue that more information is necessarily undesirable; indeed such information may be necessary to give meaning (i.e., relevancy) to a choice context even if an individual invokes an IPS that involves excluding specific attributes and even aggregating them. That is, individuals invoke procedural strategies in the form of rules that they draw on as useful devices to process information in real or hypothetical markets. Indeed aggregating does not imply that we should provide the aggregated attribute in the design, but rather that this information is often useful (it is not ignored), and a respondent prefers to be aware of it and add it up in the processing of the SC experiment. This should not be seen necessarily as cognitive burden – indeed limited information may in itself be especially burdensome where it is an incomplete representation of the attribute space that matters to an individual. The evidence suggests that aligning 'choice complexity' with the amount of information to process is misleading. *Relevancy* is what matters

KEY WORDS: ***Stated choice, information processing, relevancy, complexity.***

AUTHORS: David A Hensher

CONTACT:

Institute of Transport Studies (Sydney & Monash)
The Australian Key Centre in Transport Management, C37
The University of Sydney NSW 2006, Australia

Telephone: +61 9351 0071
Facsimile: +61 9351 0088
E-mail: itsinfo@its.usyd.edu.au
Internet: http://www.its.usyd.edu.au

DATE: May, 2004

1. Introduction

How do individuals cope with the amount of information that they are asked to evaluate in a stated choice experiment? In particular, given the number of attributes associated with each alternative and the number of alternatives in a choice set, do individuals consider all attributes (or ignore some), and do they combine attributes where feasible (e.g., in the current context, add up the travel time and/or travel cost components)? When an experienced reference alternative is used as a switch (or an appropriate memory-alternative as in case-base decision theory – see Gilboa and Schmeidler, 2002) for a stated choice (SC) design, what role might the numerical differences in attribute levels play in influencing the information processing strategy? In particular, does this influence how many attributes are excluded from the assessment (for whatever reason – relevance or cognitive burden)? What role do specific design dimensions play in how many attributes across the alternatives are ignored? These questions are illustrative of the growing array of issues that researchers need to understand if they are to make comments on the influence of the SC experiment per se as an appropriate instrument for revealing choice processes and outcomes[1].

To investigate these matters, it is necessary to design a suite of SC experiments in which we can systematically vary the dimensionality of the experiment and complement this with questions to reveal the information processing strategy (IPS) that an individual uses to assess the SC task and make a choice response. Typically, the amount of information displayed in an SC task comprises a set of attributes associated with a number of alternatives (i.e., a choice set). One of these alternatives might be a base or reference alternative described by the attribute levels experienced in real markets (in the current study, it is the trip attributes associated with a recent car commuting trip)[2]. The number of choice sets can also vary across a sample of individuals. Within each SC alternative, the number of levels of each attribute and the numerical range of each attribute might be varied, giving a choice task defined by four dimensions per choice set (i.e., number of attributes, number of alternatives, number of levels of each attribute, range of each attribute), repeated over a number of choice sets.

The design dimensionality is often referred to as the *complexity* of the SC experiment. It is typically implied that designs with more information items are more complex than those with less[3] (e.g., Arentze et al., 2003). This is potentially quite misleading since it

[1] Indeed these types of questions are associated with the information load that is faced in making any decision, and is not specific to stated choice experiments. This is all pervasive in real markets. The comment by many, that this is a specific concern for SC experiments, fails to recognise the often demanding request in revealed preference studies to provide information on the levels of attributes associated with non-chosen and/or non-experienced alternatives. This, it might be argued, is even more demanding for a respondent.

[2] The presence of a reference 'alternative' with a history of real experience is central to the case-based decision theoretic approach developed by Gilboa and Schmeidler (1995). The central tenet is that decisions are based on the relative success of actions under similar circumstances in the past. SC data has the power of richness to enable respondents to express preferences involving not only the actual memory but also *related hypothetical memories constructed from it*. These hypothetical memories consist of *cases* that combine *circumstances* that were actually encountered with various *outcomes* in addition to those actually experiences in the respective *cases*. In the language of case-based decision theory, a case is a story in which a *problem* was encountered, an *act* was chosen, and an *outcome* consequently experienced.

[3] Complexity also includes attributes that are lowly correlated, in contrast to highly correlated (the latter supporting greater ease of assessment in that one attribute represents other attributes).

suggests that complexity is an artefact of the *quantity* of information in contrast to the *relevance* of information. For example, for a specific choice outcome, an individual might feel more confident in making a considered choice response when the attributes offered are the components of travel time (e.g., free flow, slowed down and stop-start time, etc) and the components of travel cost (e.g., running cost, toll cost, etc) rather than a subset of the attributes or even an aggregated definition (i.e., total time and total cost). Knowing the numerical contribution of the toll cost in total cost may be especially useful since an individual may have a specific view on paying a toll in contrast to the outlay of normal vehicle running costs[4]. Likewise knowing the mix of free flow and congested travel time may be pertinent to the assessment, even if the aggregate time is of primary consideration[5].

The increasing amount of useful information provided can be supported by the event-splitting effect literature in economics and marketing in which 'unpacking' positive attributes of a good into multiple sub-attributes can make a good or service seem more desirable (Weber et. al., 1988, Starmer and Sugden, 1993). Conversely unpacking attribute levels that are less desirable (e.g., a greater incidence of non-free flow time in total time as revealed in the unpacking of total time) can make a good or service less desirable. Indeed the empirical evidence supports this – for a given total time, an individual generally prefers a higher proportion of free flow time. Such descriptions do matter (Hensher 2003). Even though the objective total time is the same, the outcome is different. This is referred to as the attribute-splitting effect (or the event-splitting effect when studying the sub-events). One implication is that individuals are more likely to choose particular alternatives if those containing relatively attractive consequences are sub-divided into two or more components[6]. The public policy implications are profound – the risks of ignoring relevant decomposition can be high in settings where the differential between weights is an important element of defence for investment in a specific alterative. In our empirical study this is a privately financed toll road which delivers not only travel time savings but a quality bonus in the form of a higher proportion of free flow time (See Hensher and Goodwin, 2004).

In the design of a choice experiment, the role of the reference alternative can have a substantial influence on the way in which the SC alternatives are specified and hence evaluated (see Starmer 2000 for an excellent discussion). In particular, the reference alternative can provide the memory content necessary to make the SC profiles meaningful in a preference-revelation sense. Prospect theory (Kahneman and Tversky, 1979), as an interpretation of procedural theories that assume agents draw on decision heuristics or rules when making choices, promotes the problem context as an important determinant of the choice-rule selection. One particular rule of interest herein is framing effects of which reference dependence is a popular interpretation which provides adaptive support in trading off the desire to make a good choice against the cognitive effort involved in processing the information tabled in the SC task. Case-based decision

[4] An individual register a toll each time they pay at a toll booth or pass through the electronic tolling system, whereas running costs (petrol, wear and tear) do not occur at the time of the operation, hence psychologically may be treated very differently.
[5] In toll road studies, it is common to include a quality bonus to represent the improved quality of the trip on a toll road in contrast to a free (alternative) route. It appears that this is essentially equivalent to the mix of free flow and congested travel time, which is different for a tolled and a free route.
[6] This assumes that the attribute is salient or 'relevant' to an individual.

theory promotes the central role of accumulated experience represented by a reference alternative. Starmer (2000, p 353) makes a very strong plea to support reference points:

> *"While some economist might be tempted to think that questions about how reference points are determined sound more like psychological than economic issues, recent research is showing that understanding the role of reference points may be an important step in explaining real economic behaviour in the field"*

In the current study, we pivot the levels of the attributes in the SC design off those of the overtly experienced reference alternative – a recent car commuter trip. The percentage variations around the base attribute levels used in designing a series of choice experiments are converted to actual travel times and costs for respondent assessment. This means that the attribute levels, anchored around the base, are likely to be more informative and relevant than if obtained from a less experiential setting. As a consequence of this design strategy, it is important to investigate the influence of the absolute difference between an SC design attribute level and that experienced. For example we might hypothesise that when the difference is greater, the need to retain more (or all) attributes increases so as to capture the differences with greater clarity (i.e., to make the choice easier to make)[7].

If individuals choose to ignore some attributes instead of aggregating them, then this can be explained by an editing routine under prospect theory known as the rule of cancellation which involves elimination of element (i.e., attributes) common to the prospects under consideration).

To investigate the systematic influences on the amount of information that is included out of the offered set, we propose a model in which the dependent variable is an ordered choice representing the difference between the maximum number of attributes in the design and the actual number that were maintained by the respondent in their information processing strategy (IPS) The latter is obtained by a series of questions in which respondents were asked to indicate which attributes they ignored in their assessment of the alternatives.. The zero value defines the inclusion of all attributes. Candidate influences include the dimensionality of the SC designs (detailed in the next section), the adding-up strategies, the relativity of attribute levels against the attribute profile of the experienced reference alternative and socioeconomic characteristics of the respondent. Two choice models are selected: (i) an ordered covariate heterogeneous logit model in which we can account for the systematic influences as mean effects as well as variance conditioners, and (ii) an ordered mixed logit model in which selected parameters are random to account from sources of preference heterogeneity.

The paper is organized as follows. The next section outlines the design plan of 16 separate designs that capture the main dimensionalities that are used in most empirical studies. A brief overview of the data is given followed by the set of models results for four SC settings, distinguished by the number of attributes in each choice set. The

[7] Although not investigated herein, there might be confoundment between higher absolute values when greater travel times or costs are experienced. This might suggest that the IP strategy is also dependent on the design context as well as the individual? For example, a design using SC levels pivoted around a reference alternative as a percentage difference may result in numerically greater journey times (costs), resulting in an individual preserving more attributes.

substantive implications of the analysis are set out followed by some conclusions and directions for ongoing research.

2. The Design Plan

Extant studies were used to assist in identifying candidate design dimensions (e.g., Ohler et al., 2000, White et al., 1998, DeShazo and Fermo, 2001, Dellaert et al., 1999). The five design dimensions that are varied are shown in Table 1[8].

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	2	3	2	Narrower than base
9	3	4	3	Base
12	4	5	4	Wider than base
15	----	6	----	----

The elements of the design plan are manipulated according to a master plan. The master plan has 16 runs. That is, 16 different designs are constructed to test the impact of the five design elements.[9] An example of a design is given in Appendix A. The master plan design allows for the interaction between the number of choice sets and number of alternatives as well as between the number of alternatives and number of attributes[10].

Each of the 16 designs are integrated into a Design of Designs (DoD) SC instrument each with two versions (i.e., blocking of 30 rows into sets of 15). Since these designs do not have the same number of alternatives, choice sets and attributes, and since they do not refer to the same number of attribute levels, neither do they refer to the same levels (narrow range, base range and wide range), all this is made interactive[11].
The empirical setting is a car commuter trip undertaken in the Sydney metropolitan area in late 2002. Six attributes were selected for each alternative based on previous evidence (see Hensher, in press) to characterise the options: free-flow time, slowed down time, stop/start time, variability of trip time, toll cost and running costs[12]. To explore how varying the number of attributes affects willingness to pay (WTP), the attributes were grouped according to the following patterns, noting that aggregated attributes are combinations of existing attributes[13]:

---

[8] Other possible elements might have been included but we selected those that most analysts have raised as possible sources of response bias. We excluded the ordering of attributes.
[9] As an example, suppose we have four alternatives each with six attributes at four levels in 30 runs (Two versions of 15 choice sets). The four alternatives are generic and thus the maximum number of parameters to estimate is 18 (i.e., 6×(4-1)) with 30 scenarios, which is feasible.
[10] Plus linear and quadratic effects, for each of the five dimensions.
[11] Although this might have been simpler as 16 separate CAPI's, it would have openined up the possibility of errors by the interviewer, which we wanted to minimise. The structure of the DoD SC experiment assigns the correct design, takes the correct version of it and builds the table to the dimensions of that specific design using the levels that it had specified.
[12] The trip car running cost is based on the trip distance reported by the respondent when asked about a recent trip and on average fuel consumption.
[13] This is an important point because we did not want the analysis to be confounded by extra attribute dimensions.

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

We have selected a generic design (i.e., unlabelled alternatives) for a number of reasons, including the avoidance of any confoundment with labeling. By including labeled alternatives (e.g., car, bus, train), as the number of alternatives increases, we risk confounding the effect of the number of alternatives with the labeling itself. To evaluate the effect of choice set size, we decided to use only the car as the transportation mode, increasing the number of alternatives by increasing the number of attribute bundles. In summary, making the task generic would ensure that the effect of increasing the number of alternatives is due to just that (namely, increasing that number) rather than to the labeling of the alternatives themselves. The master plan (design) gives 16 sub-designs to build as shown in Table 2[14].

[14] Each run of the design determines the specification of a choice experiment that has two versions. For example, the first row has 15 choice sets of three alternatives each presenting four attributes at three levels. For these specifications an efficient design was created. Design efficiency is an issue of great importance. Although all designs vary in efficiency, two different designs cannot be equally efficient. In a design context, design efficiency depends on the estimated parameters, thus a design that is 100% efficient is only so for a given set of parameters. Usually, the assumption is that such parameters are equal to zero. In reality, they are unlikely to be equal to zero, and so choice experiments cannot be 100% efficient. Thus, efficiency is likely to vary depending on the estimated parameters. In the current study, the real focus is on defining sensible choice tasks and so strategies such as combining specific attributes to produce fewer attributes rather than deleting attributes regarded as less important was the preferred strategy (which would reduce efficiency to some extent). Importantly, while equal 100% efficiency in an absolute sense is desirable, it is unlikely to be achievable except in special cases. Our primary focus is on classifying five task differences in order to investigate how best to trade them off in order to create a choice experiment. Discussions with Valerie Severin are appreciated.

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	3	4	3	Base
12	3	4	4	Wider than base
15	2	5	2	Wider than base
9	2	5	4	Base
6	2	3	3	Wider than base
15	2	3	4	Narrower than base
6	3	6	2	Narrower than base
9	4	3	4	Wider than base
15	4	6	4	Base
6	4	6	3	Wider than base
6	3	5	4	Narrower than base
9	4	4	2	Narrower than base
12	3	6	2	Base
12	2	3	3	Narrower than base
9	2	4	2	Base
12	4	5	3	Narrower than base

Note: Column 1 refers to the number of choice sets. The 16 rows represent the set of designs (referred to as Des0,Des1,…..,Des15 in model estimation).

The specific design pivots off of the attribute levels associated with a current car-commuting trip. As a generic design, the added alternatives are exactly the same. That is, for two design alternatives, we should not expect to find the parameter for each attribute (e.g., ‘free flow travel time’) to be different for the set of non-current alternatives. They should be the same, so we can just estimate one parameter. Therefore we do not need the attribute ‘free flow time one’ to be orthogonal to the attribute ‘free flow time two’ etc up to ‘free flow time J-1’. All we need is to make sure that the attribute ‘free flow time’ representing all non-current alternatives is perfectly[15] orthogonal to the other attributes (such as slow down time etc). This strategy reduces the whole design to a set of eight identical scenarios for each respondent. By doubling the number of scenarios we can allow some 2-way interactions. The design is made smaller because we do not need the extra orthogonality between alternatives.

The designs are computer-generated. They aim at minimizing the correlations between attributes and maximizing the amount of information captured by each choice task. An issue to take into account is finding a way that reduces the number of alternatives that are dominated by or that dominate another alternative in a choice set. In designing choice experiments, knowledge of the parameters or at least some priors (like signs) for each attribute provides a useful input. Insights from past studies determined their approximate values. A preferred choice experiment design is one that maximizes the determinant of the covariance matrix, which is itself a function of the estimated parameters[16]. The design developed herein takes into account the expected signs of the parameters (e.g., negative for the time and cost attributes). We found that in so doing,

[15] *Approximately* orthogonal is also acceptable given that some designs cannot guarantee complete orthogonality without loss of structure in terms of cognitive efficiency (in contrast to statistical efficiency).
[16] This applies to both labelled and unlabelled SC designs, although the covariance matrix is much more complex in the labelled setting.

the search eliminates dominant alternatives, which is sensible since dominant alternatives do not give any useful information if we know the signs of the parameters. The method used finds the D-optimality plan very quickly. Carlsson and Martinsson (2003) have recently shown, using Monte-Carlo simulation, that D-optimal designs, like orthogonal designs, produce unbiased parameter estimates but that the former have lower mean square errors

The levels applied to the choice task differ depending on the range of attribute levels as well as on the number of levels for each attribute. The levels are variations from the attribute value of a recent trip and are shown in Appendix B (Figures B1- B9). Extensive consideration was given to the resulting levels of attributes derived from the design, given the base levels provided by the respondent. Drawing on case-based decision theory (see Gilboa and Schmeidler, 1995), we have a strong preference for designs that pivot off of a well defined level of overt experience such that the variations in level and combination of attributes are meaningful to a respondent[17]. Without this context, there is a high risk of design effects being no more than design effects rather than genuine reflections of behavioral response.[18] The design dimensions are translated into SC screens as illustrated in Figure 1. The number of attribute levels and the range of these levels are identical within each of the 16 designs defined by the master plan. They only vary *across* designs. Each sampled commuter is given a varying number of choice sets (or scenarios), but the number of attributes and alternatives remain fixed. Variation in the number of attributes and alternatives occurs across car commuters. All analysis reported herein uses the elicitation response associated with a choice set that excludes the recent trip. Hensher (in press) contrasts the two elicitation responses.

[17] In case-based decision theory (CBDT) the idea of *meaningfulness* is equivalent to the idea of *similarity*. In CBDT, rules per se about the past are not formulated, but rather acts are evaluated by their average past performance and so any new variations (as offered in SC alternatives pivoted off the reference base) are aided by decisions criteria that can be thought of as performing implicit induction. That is they are ways to learn from past cases as to which decision should be made in a new context.
[18] Seasoned SC experiment designers will often state that the most important and challenging issue is the definition of the set of attributes and the actual numerical magnitudes to use in translating the design code levels into numerical values that mean something to a respondent. This task is more an art than a science.

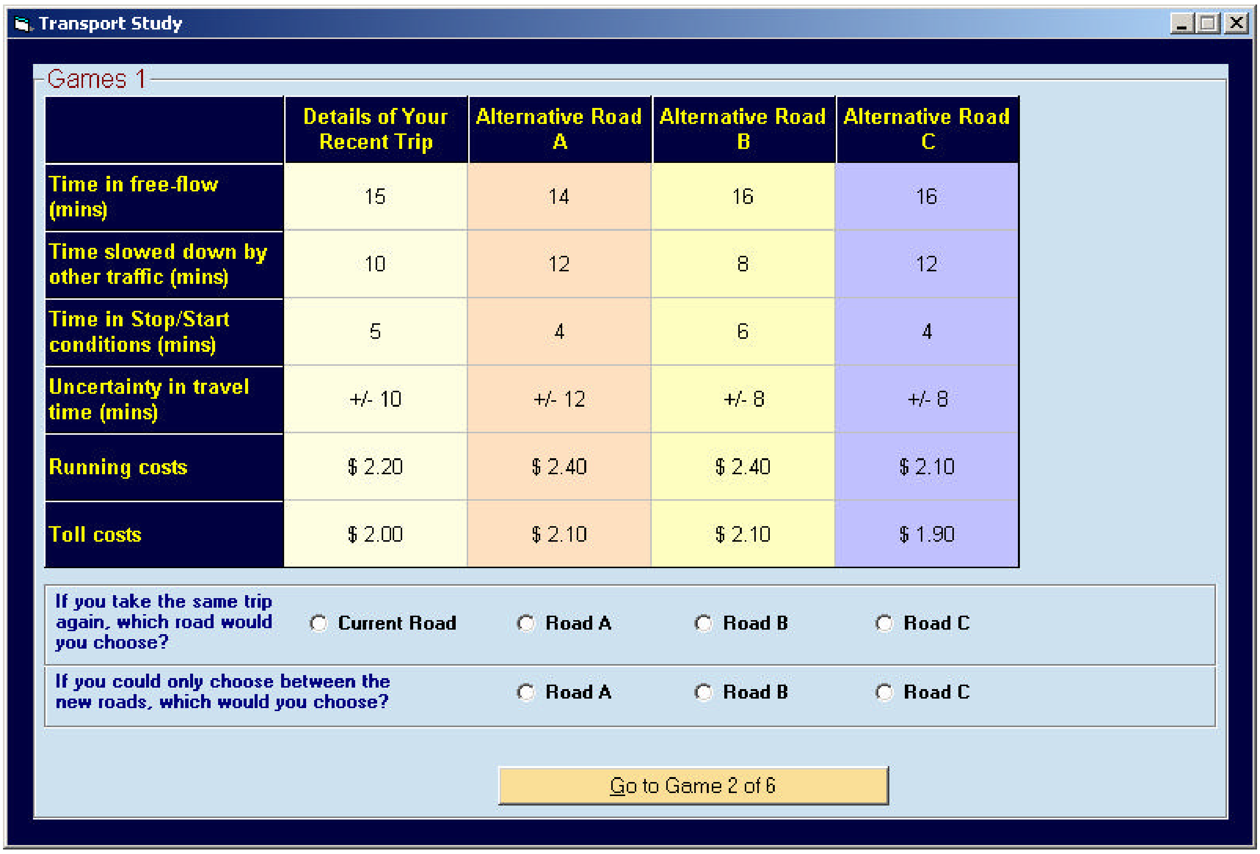

Figure1. An example of a stated choice screen

3. Modeling Approach

Separate ordered heterogeneous logit (OHL) and ordered mixed logit (OML)[19] models have been selected to estimate models for the design subsets associated with each of four numbers of attributes associated with each alternative (i.e., 6, 5, 4 and 3). This is necessary in order to define the dependent variable as the number of attributes out of the full set that were included in the information processing leading to a choice response[20]. Treating this (ranking) scale *as if it were both continuous and an interval scale* and applying ordinary least squares regression to analyze the relationship between the amount of information processed and the explanatory variables is not a statistically valid approach (Winship and Mare, 1984).

The ordered logit model allows one to include ordinal dependent variables into the choice model in a way that explicitly recognizes their ordinality and avoids arbitrary assumptions about their scale. It defines points on the *observed scale* as thresholds. The essence of the approach is an assumed probability distribution of the continuous variable that underlies the observed ordinal dependent variable.
Formally, let Y^* denote an unobserved (or latent) continuous variable (-8 < Y^* < +8), and μ_0 $\mu_?$,..., μ_{J-1}, μ_J denote the cut-off or threshold points in the distribution of Y^*,

[19] Mixed logit models are well documented in a growing number of papers and books. See Train (2003) and Hensher et al. (2004) for overviews.
[20] In our empirical study this makes good sense because an individual only saw a fixed number of attributes across the set of choice sets.

where μ_0= -8 and μ_J = +8 . Define Y to be an ordinal (observed) variable such that Y = j iff $\mu_{J-1} = Y^* = \mu_J$; j = 1,2,...,3. Since Y^* is not observed but Y is observed, its mean and variance are unknown. Statistical assumptions must be introduced such that Y^* has a mean of zero and a variance of one. To operationalise the model, we need to define a relationship between Y^* and Y.

The ordered choice model is based on the following specification: There is a latent regression equation (1).

$$Y_i^* = b¢\mathbf{x}_i + \varepsilon_i,\ \ \varepsilon_i \sim F(\varepsilon_i\,|q),\ E(\varepsilon_i) = 0,\ \mathrm{Var}(\varepsilon_i) = 1 \qquad (1)$$

The observation mechanism results from a complete censoring of the latent dependent variable as follows:

$$\begin{aligned} Y_i \quad &= 0 \text{ if } Y_i \le \mu_0, \\ &= 1 \text{ if } \mu_0 < Y_i \le \mu_1, \\ &= 2 \text{ if } \mu_1 < Y_i \le \mu_2, \\ &\ldots \\ &= J \text{ if } Y_i > \mu_{J-1}. \end{aligned} \qquad (2)$$

The probabilities which enter the log likelihood function are given by equations (3) and (4).

$$\mathrm{Prob}(Y_i = j) = \mathrm{Prob}(Y_i^* \text{ is in the } j\text{th range}) \qquad (3)$$

$$\mathrm{Prob}(Y_i = j) = F((\mu_j - b¢\mathbf{x}_i)) - F((\mu_{j-1} - b¢\mathbf{x}_i)),\ j = 0,1,...,J \qquad (4)$$

We can generalize the unobserved variance to accommodate multiplicative heterogeneity, as we do herein, and make this variance a function of contextual variables (as in OHL) or introduce preference heterogeneity via a set of random parameters as in OML.

A direct interpretation of the parameter estimates from an OHL or OML model is not informative given the logit transformation of the choice dependent variable required for model estimation. We therefore provide the marginal effects which have substantive behavioural meaning, defined as the derivatives of the choice probabilities (Hensher et al. 2004). A marginal effect is the influence a one unit change in an explanatory variable has on the probability of selecting a particular outcome, *ceteris paribus*[21]. The marginal effects need not have the same sign as the model parameters. Hence the statistical significance of an estimated parameter does not imply the same significance for the marginal effect (see equation (5)).

$$\partial\mathrm{Prob}(y_q = j)/\partial\mathbf{x} = P_j(b_j - \bar{b}),\ \ \bar{b} = \Sigma_j P_j b_j. \text{ (defined below as } d_j\text{)} \qquad (5)$$

[21] This holds for continuous variables only. For dummy (1,0) variables, the marginal effects are the derivatives of the probabilities given a change in the level of the dummy variable and thus represent the influence of a change in level of the variable upon the probability of choosing a given outcome, *ceteris paribus*.

Neither the sign nor the magnitude of d_j need bear any relationship to those of b_j. $b_{\mathbf{j}}$, equal to $\partial \log(P_j/P_0)/\partial \mathbf{x}$, is commonly defined as an interpretation of the model parameters. Where the estimated parameters are random, as in mixed logit, the marginal effects are more complex. Since there is no meaningful conditional mean function to manipulate, we consider, instead, the effects of changes in the covariates on the cell probabilities. These are:

$$\partial \mathrm{Prob}(\mathrm{cell}\ j)/\partial \mathbf{x}_i \;=\; (f(\mu_{j-1} - \mathrm{b}'\mathbf{x}_i) - \mathrm{f}(\mu_j - \mathrm{b}'\mathbf{x}_i)) \times \mathrm{b}$$

where $f(\mu_{j-1} - \mathrm{b}'\mathbf{x}_i)$ is the logistic density, $\Lambda(\mu_{j-1} - \mathrm{b}'\mathbf{x}_i)(1-\Lambda(\mu_{j-1} - \mathrm{b}'\mathbf{x}_i)$.

Potential influences assessed in the ordered logit model are:

1. The dimensionality of each SC task (i.e., number of levels of each attribute, numerical range of these levels, the number of alternatives);
2. The number of choice sets to be evaluated;
3. The deviation of the design attribute levels from the reference (or overtly experienced) alternative;
4. The use of ‘adding up’ attributes where this is feasible (e.g., travel time components); and
5. The socioeconomics characteristics of the respondent.

4. Empirical Analysis

Computer aided personal interview (CAPI) surveys, 514 in total, were completed in the Sydney metropolitan area between 19 October and 23 November 2002[22]. Sampling was stratified random according to the residential location of the household. Screening questions established eligibility in respect of commuting by car. Quotas were imposed for three trip lengths: less than 30 minutes (256), 30-60 minutes (190) and 60-90 minutes (60). Further details are given in Hensher (in press). Final OHL and OML models for each of the four settings are given in Tables 3 and 5 with their respective marginal effects in Tables 4 and 6[23].

[22] This was preceded by a pilot survey of 36 commuters, which was sufficiently large (after expansion of choice sets) to enable estimation of multinomial logit models to at least assess the parameter estimates in respect of sign and relative magnitude (on marginal effects).
[23] We have retained all explanatory variables from the various sources of influence on IP strategy that has a t-value greater than 0.10.

Table 5 Ordered Heterogeneous Logit Models for IPS models with 6, 5, 4 and 3 attributes per alternative.

Attribute	Units	OHL6		OHL5		OHL4		OHL3	
Constant		4.2065 (4.78)		2.688 (5.31)		11.743 (7.0)		-.0042 (-.93)	
Design Dimensions:									
No. of levels	Number	0.5299 (4.22)		0.0427 (0.98)		-.1383 (-.61)		.0009 (.82)	
Narrow attribute range	1,0	1.2518 (6.03)		1.2711 (4.43)		2.2922 (2.84)		-.00004 (-.08)	
No. of alternatives	Number	-0.8956 (-3.72)		-.8163 (-5.23)		-2.1721 (-4.62)		.00013 (0.6)	
Framing around Base Alt:									
Free flow time for Base (total time for OHL3) – SC alternative	Minutes	0.0146 (1.96)		.0319 (5.07)		-		-.0000027 (-.5)	
Congested time for Base (total time for OHL3) – SC alternative	Minutes	0.0142 (2.42)		-.0083 (-1.91)		.0354 (3.43)		N/A	
Attribute Packaging:									
Adding travel time components	1.0	-0.4922 (-3.49)		-.9054 (-7.56)		1.2245 (4.19)		N/A	
Variance Covariates:									
No. of choice sets	Number	0.0345 (6.41)				.0615 (8.23)		-.2877 (-4.1)	
Free flow time for Base (total time for OHL3)– SC alternative	Minutes	-		-.0187 (-3.27)		-.0049 (-1.60)		.00074 (.12)	
Congested time for Base (total time for OHL3) – SC alternative	Minutes	0.0093 (4.58)		-		.0072 (2.79)		-	
Personal income	$'000s	0.0024 (3.40)		-.0035 (-2.0)		.0054 (7.78)		-.0406 (-8.4)	
Threshold Parameters:									
Mu1		1.7896 (12.9)		2.697 (8.72)		7.435 (8.91)		.05568 (1.47)	
Mu2		3.9928 (13.6)		-		13.854 (8.74)		-	
Count of Choice Responses:									
		*	obs	*	obs	*	obs	*	obs
0		6-0	1014	5-0	1415	4-0	810	3-0	1698
1		6-1	876	5-1	1080	4-1	2214	3-1	1212
2		6-2	900	5-2	66	4-2	684	3-2	48
3		6-3, 6-4	585			4-3	99		
Log-Likelihood		-4518.45		-1880.52		-3847.70		-2077.95	

Notes: * defines max # attributes minus #ignored.

Table 6 Marginal Effects Derived from Ordered heterogeneous Logit Models for IPS models with 6,5,4 and 3 attributes per alternative.

Attribute	OHL6				OHL5		
	0	1	2	3	0	1	2
Design Dimensions:							
No. of levels	-.068	-.008	.023	.029	-.014	.013	.001
Narrow attribute range	-.162	-.025	.073	.091	-.421	.389	.0320
No. of alternatives	.116	.016	-.046	-.057	.271	-.250	-.0205
Framing around Base Alt:							
Free flow time for Base – SC alternative	-.0019	-.0003	-.0009	-.0013	-.011	.010	.0008
Congested time for Base – SC alternative	-.0018	-.0003	.0009	.0012	.003	-.003	-.0002
Attribute Packaging:							
Adding travel time components	.064	.018	-.052	-.065	.300	-.277	-.023
Variance Covariates:							
No. of choice sets	.096	-.083	-.057	.077			
Free flow time for Base – SC alternative					-.193	-.004	.1970
Congested time for Base – SC alternative	.764	-1.01	-.691	.940			
Personal income	.227	-.258	-.178	.241	-.003	-.00006	.0031

Attribute	OHL4				OHL3		
	0	1	2	3	0	1	2
Design Dimensions:							
No. of levels	.008	.00001	-.007	-.001	-.067	.067	Very small
Narrow attribute range	-.137	-.0002	.118	.0189	.003	-.003	Very small
No. of alternatives	.129	.0002	-.112	-.0179	-.009	.009	Very small
Free flow time for Base (total time for OHL3) – SC alternative					.0002	-.0002	Very small
Congested time for Base – SC alternative	-.002	.000001	.002	.0003			
Adding travel time components	-.073	-.0001	.063	.0101	-	-	
No. of choice sets	.508	-1.016	.316	.192	-.00002	.00002	Very small
Free flow time for Base (total time for OHL3) – SC alternative	2.60	-5.20	1.618	.9829	.00008	-.00008	Very small
Congested time for Base – SC alternative	-.031	.061	-.019	-.012			
Personal income	-.481	.963	-.299	-.182	.0000	.0000	Very small

Table 7 Ordered Mixed Logit Models for IPS models with 6,5,4 and 3 attributes per alternative. RPL = random parameter, FP = fixed parameter

Attribute		OML6 (0-3)	OML5	OML4	OML3
	Units				
Constant		4.0372 (4.75)	7.494 (5.51)	5.1041 (8.44)	-2.244 (-4.24)
Non-Random Parameters:					
No. of choice sets	Number	-	-	-	-.3305 (-5.1)
Number of levels	Number	RPL	RPL	RPL	2.7935 (5.84)
Narrow attribute range	1,0	1.1129 (6.60)	2.4602 (3.91)	.7829 (2.16)	-
No. of alternatives	Number	-.8036 (-3.27)	-1.7753 (-5.04)	RPL	RPL
Adding travel time components	1.0	-.3951 (-3.10)	-1.1360 (-5.3)	.5976 (4.05)	N/A
Free flow time for Base (total time for OML3)– SC alternative	Minutes	0.0173 (2.19)	0.0262 (3.10)	-.0068 (-.76)	-.0089 (-1.43)
Congested time for Base (total time for OML3) – SC alternative	Minutes	0.0140 (2.08)	-.0207 (-2.50)	.0121 (1.83)	N/A
Personal income	$000s	-.0052 (-3.13)	-.01873 (-5.3)	-.0052 (-3.63)	-.0110 (-4.5)
Random Parameters:					
No. of levels	Number	0.4824 (3.64)	-.1645 (-1.80)	-.0851 (-.74)	FP
No. of alternatives	Number	FP	FP	-.7948 (-4.0)	-1.1426 (-4.8)
Scale Parameters					
No. of levels	Number (normal)	0.5539 (7.22)	0.4508 (4.15)	.4904 (7.56)	FP
No. of alternatives	Number (normal)	FP	FP	.0604 (.16)	.5985 (4.0)
Threshold Parameters:					
Mu1		1.7088 (13.60)	4.6264 (10.60)	3.762 (15.33)	6.2020 (6.47)
Mu2		3.6897 (14.16)	-	6.548 (16.21)	-
Log-Likelihood		-4523.64	-1834.76	-3926.26	-2136.92

Table 8 Marginal Effects Derived from Ordered Mixed Logit Models for IPS models with 6, 5, 4 and 3 attributes per alternative.

Attribute	OML6				OML5		
	0	1	2	3	0	1	2
No. of levels	-.0814 (-3.76)	-.0343 (-2.55)	0.078 (1.61))	.0370 (.85)	.0402 (1.9)	-.039 (-1.8)	-.0012 (-.97)
Narrow attribute range	-.143 (-4.87)	-.128 (-20.8)	.1478 (0.53)	0.123 (5.85)	-.547 (-59.7)	.520 (44.6)	.0271 (7.57)
No. of alternatives	0.136 (3.38)	.0571 (2.38)	-.131 (-1.42)	-.062 (-0.92)	.433 (5.1)	-.421 (-5.1)	-.013 (-.86)
Adding travel time components	0.0919 (11.2)	.0538 (3.45)	-.063 (-0.90)	-.034 (-.33)	.276 (20.5)	-.264 (-21.8)	-.0121 (-.50)
Free flow time for Base – SC alternative	-.0029 (-2.23)	-.0012 (-1.90)	0.003 (1.70)	0.001 (0.67)	-.006 (-3.1)	.006 (3.1)	.0002 (.77)
Congested time for Base – SC alternative	-.0024 (-2.10)	-.0010 (-1.90)	0.002 (1.72)	0.001 (0.66)	0.005 (2.5)	-.005 (-2.5)	-.0001 (-.74)
Personal income	0.0009 (3.26)	0.0004 (2.28)	-.0008 (-1.78)	-.0004 (-.74)	.005 (7.7)	-.004 (-7.7)	-.0001 (-.97)
Attribute	OML4				OML3		
	0	1	2	3	0	1	2
Number of choice sets					.077 (5.7)	-.077 (-5.6)	-.0003 (-1.8)
No. of levels	.010 (.74)	-.0004 (-.20)	-.0088 (-.66)	-.0008 (-.47)	-.650 (-6.5)	.647 (6.5)	.003 (1.7)
Narrow attribute range	-.077 (-3.2)	-.029 (-2.28)	.097 (.1.6)	.0049 (6.8)			
No. of alternatives	.093 (4.0)	-.0004 (-.20)	-.0823 (-5.5)	-.007 (-1.1)	.266 (5.3)	-.265 (-5.3)	-.0013 (-1.8)
Adding travel time components	-.080 (-2.9)	.0219 (1.3)	.054 (1.3)	.0045 (4.8)			
Free flow time for Base (total time for OML3) – SC alternative	.0008 (.76)	-.00003 (-.20)	-.0007 (-.71)	-.00006 (-.53)	.002 (1.5)	-.002 (-1.43)	-.00001 (-1.1)
Congested time for Base – SC alternative	-.0014 (-1.82)	.00006 (.21)	.0012 (1.73)	.0001 (.85)			
Personal income	.0006 (3.9)	.000024 (-.22)	-.0005 (-2.5)	-.00005 (-.88)	.003 (4.7)	-.2646 (-5.3)	-.00001 (-1.9)

The evidence identifies a number of statistically significant influences on the amount of information processed, given the maximum amount of information provided. Individuals clearly self-select information to process in stated choice studies, just as they do in real markets where the (transaction) costs of seeking out, compiling and assessing large amounts of (potentially useful) information is often seen as burdensome and/or as not producing sufficient benefits. While the evidence herein cannot establish whether an information reduction strategy is *strictly* linked to behavioral relevance or to a coping strategy for handling cognitive burden, both of which are legitimate paradigms in real markets, it does provide important signposts on how information provided within a specific context is processed to reflect what we broadly call the relevancy paradigm. Something is relevant either because it does influence a choice in a real sense of inherent preference and/or in discounting its potential role as a coping mechanism. We do this daily in most decisions we make and hence it could be argued that this information processing strategy is not unique to stated choice studies, but a commonly practiced IP strategy.

Taking a closer look at each OHL and OML model, there are some important empirical outcomes. We discuss the evidence under the following headings: (i) design dimensions, (ii) framing around the base, (iii) attribute packaging, (iv) variance decomposition, and (v) socio-economic effects.

4.1 Design Dimensions

Beginning with the six-attributes per alternative models (OHL6, OML6), the statistical significance of the three dimensions of a choice set (namely number of alternatives, number of levels per attribute and range of an attribute (narrow versus not narrow)) is high. The direction of influence is shown in the marginal effects for each level of the dependent variable. For the number of levels per attribute and attribute range, the negative marginal effects for 0 and 1 levels of the response variable and the positive marginal effects for levels 2 and 3 suggests that the probability of preserving more (or all in case of the zero response) attributes from the offered set increases *dramatically* as the number of levels per attribute declines and the attribute level range widens. What this may indicate is that *if each attribute across the alternatives in a choice set provides less variability over a wider range (in sense of proximity and variability within a range), then it is useful to preserve the information content of all attributes as necessary to assist in differentiation*. Furthermore, as we increase the number of alternatives to evaluate the importance of maintaining more (including all) attributes increases, again as a possible mechanism for ensuring greater clarity of differentiation between the alternatives.

The evidence for the narrow attribute range is reinforced very strongly for OHL5/OML5 and OHL4/OML4; however the attribute range is statistically non-significant for OHL3/OML3. The range has the greatest marginal effect for five-attributes (which differs from six-attributes in the presentation of a single travel cost attribute). This adds further support to a view that a narrower attribute range tends to decrease the probability of preserving all or most attributes regardless of the available attribute set, in contrast to a wider range where attribute preservation is increasingly more relevant to assist in making a choice. Another way of stating this is that if an analyst continues to include, in model estimation, an attribute across the entire sample that is not marked for preservation, then there is a much greater likelihood of biased parameter estimates in circumstances where the attribute range is narrower than wider[24].

For the number of alternatives, statistical significance and maintained signs prevail for all OML and OHL models except OHL3. This reinforces the view above that greater differentiation within the attribute set (levels and range) is preferred as the number of alternatives to evaluate increases, which is secured by preserving all attributes[25]. This is an important finding that runs counter to some views, for example, that individuals will tend to ignore increasing amounts of attribute information as the number of alternatives increases. *Our evidence suggests that the IP strategy is dependent on the nature (i.e., profile) of the attribute information and not strictly on the quantity.*

[24] This has interesting implications for the growing evidence that WTP for an attribute tends to be greater under a wider range for the numerator attribute. Simply put, the greater relevance in preserving the attribute content under a wider range will mean that such an attribute is relatively more important to the outcome than it is under a narrow range specification, and hence a higher WTP is inferred.

[25] Up to the number evaluated in this study. In studies with a greater number of attributes, there be a threshold beyond which additional attributes are not preserved.

The influence of the number of levels falls away for OHL5/4/3 and OML5/4 (as a random or fixed parameter) but retains its statistical significance in OML3 as a fixed parameter. The marginal effect for OML3 retains its negative sign under maintenance of all attributes but becomes positive and significant when ignoring one attribute. The evidence on preference heterogeneity engendered with more levels of an attribute for the larger set of attributes raises speculation about the increase in such heterogeneity beyond six attributes. Being well within the magic number seven for each alternative (Starmer, 2000), we might speculate about high levels of unobserved heterogeneity as we increase the number of levels per attribute for an increasing number of attributes to evaluate per alternative. This has intuitive appeal. At the other end of the attribute number spectrum, the evidence that the number of levels is statistically significant for the three-attribute model as a fixed parameter, suggests much lower (true or spurious?) preference heterogeneity at the lower end of the spectrum of attributes per alternative. *That is, through eliminating information that may be relevant by limiting the number of attributes and level per attribute to process, we may be removing the opportunity to reveal true preference heterogeneity (and a consequent spurious support for a fixed parameter result).*

Overall, we see a picture emerging that design dimensionality seems to have less of an influence on the IP strategy when we have fewer items to process. This makes good sense but should not be taken to imply that designs with fewer items are preferred, but that preference heterogeneity in invoking an IP strategy appears to decline substantially as the information content declines, for real or spurious reasons. Contrariwise, individuals appear to increasingly invoke a relevancy strategy as the amount of information to process increases. The need to capture this growing heterogeneity in IP strategies is clear and may be captured through the inclusion of an IPS 'selectivity correction'[26] variable in the behavioral choice models.

[26] A selectivity correction variable is derived from a recognition that, within the setting of a specific stated choice experiment, there can exist a range of IP strategies from which one may be selected by a sampled respondent, to assist them in evaluating a specific SC scenario. This selection process, if correlated with the choice process, should be accommodated and explicitly incorporated into the stated choice model. The statistical significance of the parameter linking this IPS index is a way of correcting for the influence of different IP strategies in processing SC experiments. The IPS index can take a number of forms. An IPS Selectivity correction involves recognition of the potential correlation between the unobserved components of the behavioural and IPS choices and the determination of a method for handling the influence of the IPS on behavioural choice response. Given :

$$E(\eta) = \underset{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_J}{E} [E(\eta|\delta_i = 1)|\ \varepsilon_1,\varepsilon_2,\ldots,\varepsilon_J]$$

and the assumption that the ε_i 's are distributed extreme value type I;

i.e., $F_\varepsilon(\varepsilon_1, \varepsilon_2, \ldots \varepsilon_J) = \exp\ [-\Sigma\ \exp\ (-\varepsilon_i / \mu)]$ with scale parameter $\mu > 0$; the selectivity correction formula is:

$$E(\eta|\ \delta_i = 1) = -(\frac{6}{\pi^2})\ \rho_i\ \sigma\ (\frac{I-1}{J}(\text{Log}\ \hat{P}_i) + \sum_{\substack{j=1 \\ j\neq i}}^{J} (\frac{\text{Log}\ \hat{P}_j}{J})\ (\frac{\hat{P}_j}{1-\hat{P}_j}))$$

The coefficient of this *selectivity correction variable is* $-(6/\pi^2)\ \rho_i\ \sigma$, where σ is the standard error of the estimate and ρ_i is the correlation between the error terms of the two models. Given the estimated parameter for selectivity correction, ρ_i can be derived.

4.2 Framing around the Base

The theoretical argument promoted in prospect theory and in case-based decision theory for reference points (see Starmer, 2000 and above) is supported by our empirical evidence. Two variables represent the framing around an experienced alternative as a context that is not biasing, but essential to the realism of the decision process. We have framed the level of each attribute relative to that of the experienced recent car commute as (i) free flow time for current (or base) minus the level associated with an attribute and alternative in the SC design and (ii) the congested (or non-free flow) travel time for the base minus the level associated with each SC alternative's attribute. Where the travel time attribute is total time, we have used the single framing attribute (i.e., OHL3/OML3).

The evidence in OHL6 and OML6 indicates that reference dependence is a significant influence on the IP strategy. When the difference in attribute magnitude increases, the probability of including more attributes in the selection process decreases across all IP response levels for free flow, but changes direction for congested time for all IP response levels that remove attributes[27].

For free flow time, this supports the role played by all (or most) attributes in narrowing down the choices, but importantly highlights how much easier it is to process information where the relativities are much greater. As the attribute magnitudes move closer, individuals appear to use some approximation paradigm in which closeness suggests similarity and hence ease of eliminating specific attributes[28] because their role is limiting in differentiation. This contrast can draw on regret theory (Loomes and Sugden, 1987) in which large differences between what you get from a chosen alternative and what you might have obtained from an alternative, give rise to disproportionately large regrets leading to individuals preferring greater certainty in the distribution of regret by choosing the alternative they have experience with. This is the same as staying with the 'safe bet'.

The relative congestion time framing effect changed sign in both OHL5 and OML5 for the IP response 0, resulting in an increase in the probability of selecting an IP strategy in which all attributes are preserved. The reversal effect is curious given that it arises where the reference to tolls has been removed. Paying a toll is often seen as the price to eliminate levels of congested traffic. In the presence of a toll, car commuters typically perceive both a time saving benefit and a re-mixing of the free flow and congested time in favor of free flow time, where free flow time has a lower marginal value than congested time. When offering a total cost for a trip (with no contextual explanation of its composition) which is one less attribute, we see strong support for preserving all attributes, suggesting (maybe) that in the specific setting under review, that adding up the two cost attributes is sensible before evaluating the alternatives[29], giving the same information load as exhibited in OHL6. That is, when the two cost components are

[27] For congested time the sign change for the marginal effects is not statistically significant (as reported for OML6).
[28] This result reinforces the evidence on the influence of the number of levels and range of attributes in the SC design.
[29] Aggregation is a particular form of preservation strategy (in contrast to the exclusion or cancellation strategy).

aggregated before assessment, the response level 1 for OHL6 is equivalent to the response level 0 for OHL5. This may suggest that if particular sub-sets of attributes are best aggregated before assessment, then one might not want to include the individual attributes in the design as separate effects.

Reference dependency not only has a direct (mean) influence on the IP strategy; it also plays an important role in accounting for heterogeneity in the variance of the unobserved effects in the OHL models. For example, the framing of congested time in OHL6 and OHL4 has a strong statistical influence on the variance. The marginal effects indicate that an increase in the gap between the base and SC attribute levels for congested time, which increases the variance of the unobserved effects in OHL6 and OHL4, leads to an increase in the probability of an individual adopting an IP strategy for OHL6 and a decrease for OHL4, in which all attributes are preserved. The directional impact is not the same in all cases for reference dependency through the parameterization of the mean and variance, but significantly we observe that when a particular reference dependency is statistically significant for the mean (in the case of free flow and congestion time) it carries this significance into the variance. That is we are seeing the influence of framing via two impact points, with the effects mostly in the same direction (ie mutually reinforcing). Thus one might suggest that the failure to account for the influence of reference dependency on the variance of the unobserved effects will under-estimate the probability of selecting a specific IP strategy.

4.3 Attribute Packaging

The event-accumulation rule in stage 1 editing under prospect theory is consistently strong, for the aggregation of travel time components, across all models (excluding the 3-attribute model where travel time is a single attribute)[30]. The mean parameter estimate is negative for six and five attributes and positive for four attributes, producing positive marginal effects for the 0 response level for the six and five attribute and negative marginal effects for the four-attribute model.

For the six-attribute model, the statistically significant positive marginal effect applies to response levels 0 and 1, which suggests that if a respondent wishes to preserve the information content of all or most attributes, then the probability of processing them through an aggregation rule increases. The effect is strongest when no attributes are excluded. Indeed it appears from other research (see Hensher, 2003) that evaluating components and aggregating them is not strictly equivalent to adding up attributes and then evaluating the aggregated attribute.

A similar set of findings apply to the five-attribute design except that the positive marginal effect is much stronger than for six attributes, with the consequence of a sign change for response level 1. Given the point in section 4.3 that the difference between the five and six attribute models is the aggregation of cost (running plus toll cost)[31], then it is not surprising to see the aggregation of time effect coming through much stronger

[30] Hensher (in press) presents the data set in detail. We find that over 80% of the sample who evaluated the six, five and four attribute designs added up components of travel time.

[31] For the six-attribute design, 68% of the sample added up the two cost attributes, equivalencing the five and six-attribute designs in a quantity sense.

under response level 0 for the five-attribute model, since it appears to be equivalent to the six-attribute model's combined response levels 0 and 1. The aggregation of cost versus preserving its components was not found to have a significant influence as an information-processing packaging strategy.

The four-attribute model produces the opposite directional impact, suggesting that when the travel time components are added up (in this case only two attributes: free flow and the pre-aggregated components of congestion time), the probability of preserving all but one attribute decreases. Checking the data shows that the attribute removed was trip time variability, reducing the assessment to a comparison on total time and total cost. This may suggest that when one gets down to so few attributes, there is a sense of simplicity being imposed on the respondent from an SC design that has limited information to process, which can maintain its relevancy through aggregation as a simple assessment of two attributes.

There is clear evidence that a relevant simplification rule is re-packaging of the attribute set where possible through addition. This is not a cancellation strategy but a rational way of processing the information content of component attributes and then weighting this information (in some unobserved way) in comparing alternatives.

4.4 Choice set effects and Variance decomposition in OHL

Under the heterogeneous logit specification (OHL) we distinguish between the influence of a design or process attribute as a mean effect, and its conditioning of the variance of unobserved effects, the latter being the mechanism for revealing the presence of preference heterogeneity in the IP strategy[32]. The most interesting result is the role of the number of choice sets in variance decomposition. Figures 2-5 graph the distribution of respondents across the IP strategies for each number of evaluated choice sets. Beginning with the IP strategy where the maximum number of attributes are not preserved, (i.e., four, two, three and two respectively for six, five, four, and three attributes per alternative), we have very few respondents. As an individual preserves an increasing number of attributes in the IP strategy (up to the full set), we find substantial differences in the incidence of IP strategy selection within each SC design, both within and between the number of choice sets except for the four-attribute case.

[32] In contrast the ordered mixed logit model uses random parameterization of specific attributes to reveal preference heterogeneity

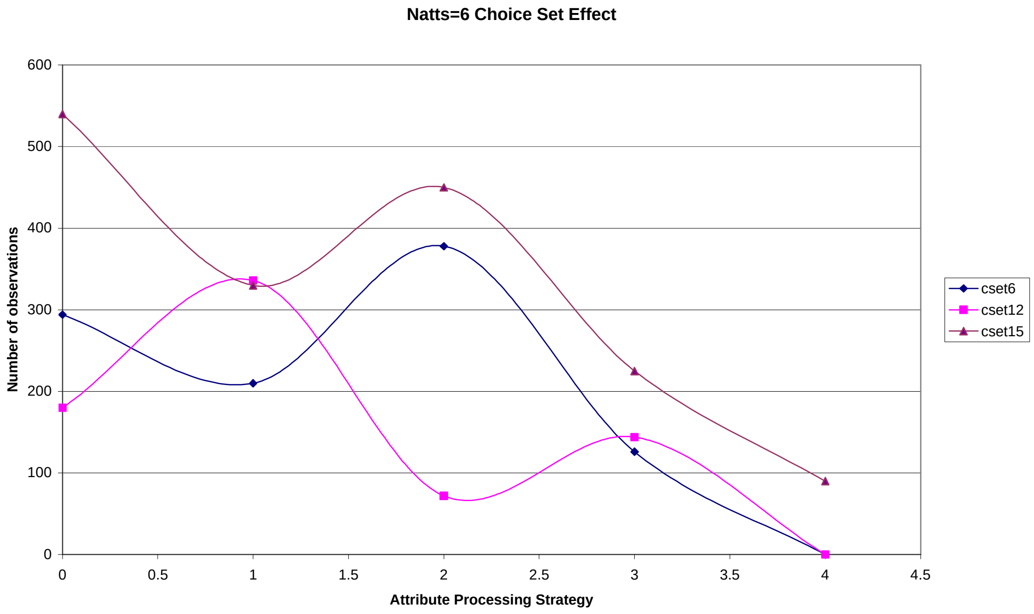

Figure 2 Impact of Number of Choice Sets on Information Processing Strategy for six attributes

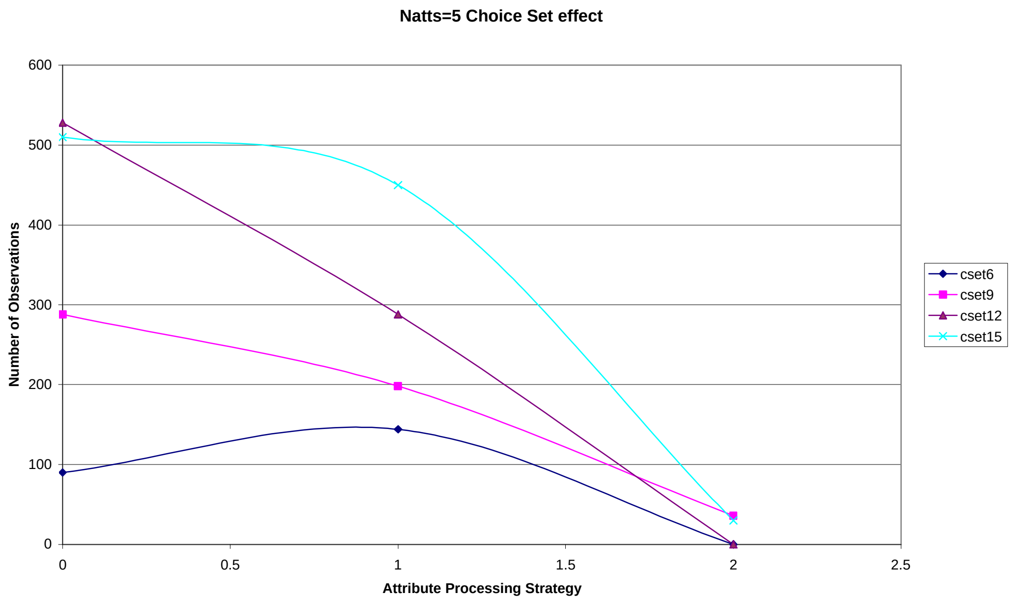

Figure 3 Impact of Number of Choice Sets on Information Processing Strategy for five attributes

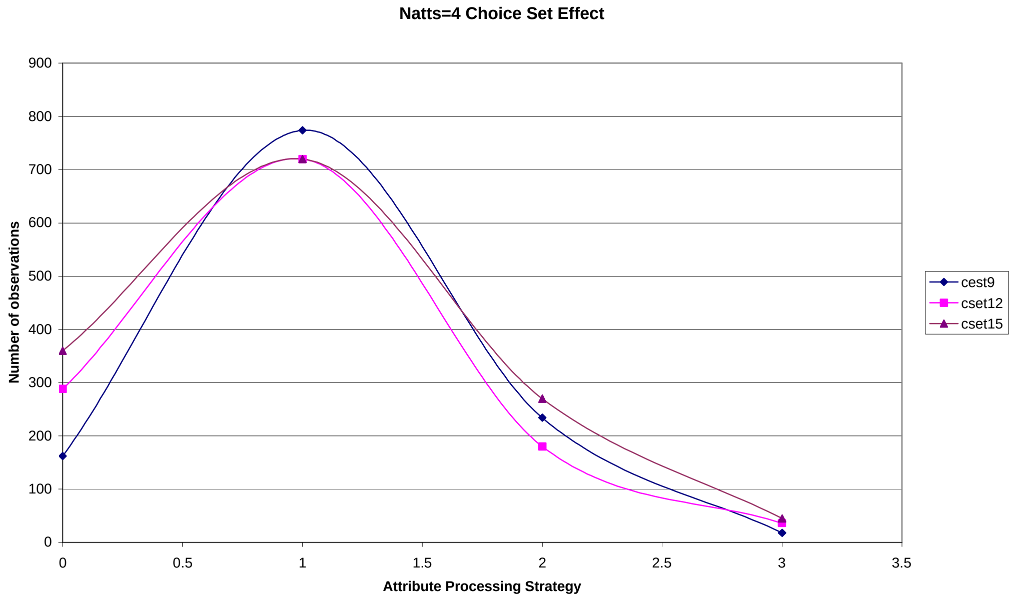

Figure 4 Impact of Number of Choice Sets on Information Processing Strategy for four attributes

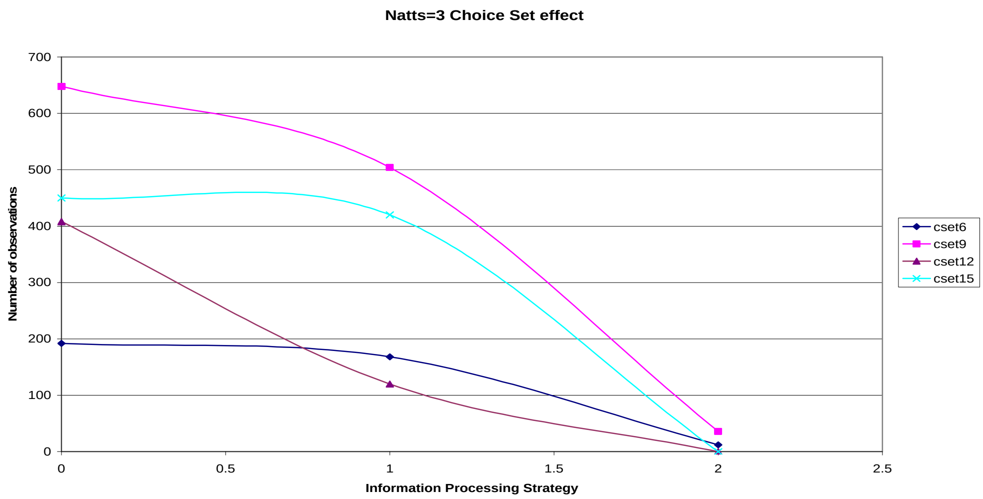

Figure 5 Impact of Number of Choice Sets on Information Processing Strategy for three attributes

The empirical results in Table 5 support the descriptive evidence of a systematic variation in the number of choice sets on the IP response, and indeed the marginal effects are the greatest. However, we also find evidence of a systematic influence of the number of choice sets in OHL6 and OHL5; albeit with very much smaller marginal effects. While the number of choice sets evaluated has no direct influence on processing *each choice set per se*, the evidence from the OHL models suggests that there is an accumulation process effect being identified when we pool the data on each choice set, and that its impact is primarily on the variance of the unobserved effects and not via the mean.

For OHL6, as the number of choice sets increases from six to 15, (and for OHL4 as the number increases form nine to 15), the variance of the unobserved influences increases as a reflection of the increasing amount of accumulated information that is being processed. The variance decreases in the case of OHL3 (with the number of choice sets varying from six to 15). This suggests some degree of interdependence between the choice sets in respect of IP response[33]. True independence would suggest that the number of choice sets has no impact on the IP strategy. Except for OHL4, however, the marginal effects are very small, especially for OHL3.

When we introduce the number of choice sets into the mixed logit model (OML6, OML5, OML4) through random or fixed parameterization, it has no statistically significant influence (indeed its t-value is extremely low), suggesting that preference heterogeneity in terms of IP strategy cannot be linked to the number of choice sets per se. This is not the case for three-attributes (OHL3) where the statistically significant fixed parameter effect supports the view that increasing the number of choice sets reduces the probability of preserving all attributes.

The evidence herein, while not systematically the same across all four SC designs, provides accumulating evidence to support the suggestion that the number of choice sets is more a matter of cognitive recognition of other possible influences (resident in the variance term capturing the unobserved effects) on how individuals structure an IP strategy as they progresses through a series of choice tasks.

4.5 Socio-economic Impacts

The respondent's personal income has a statistically significant influence on the IP strategy in all OHL and OML models. The marginal effects for the 0 response level are positive and significant in all mixed logit models, indicating that the probability of preserving all attributes increases as an individual's personal income decreases. In the ordered heterogeneous logit models, personal income enters into the variance decomposition; however despite its statistical significance as a mean parameter estimate, the transformation into a marginal effect produces some very small (albeit negligible) impacts for OHL5 and OHL3.

The positive marginal effect for preservation of all attributes in the six-attribute model suggests that a higher personal income tends to reduce the variance of the unobserved effects, resulting in a higher probability of preserving the full set of attributes. This is the opposite of the finding when personal income enters as a mean effect directly on the IP strategy. Clearly for six attributes, we have an effect working in different directions according to whether its impact is via the mean or the variance. For the three-attribute model we also have opposite effects but this time it is the reverse of the six-attribute case.

[33] We could not find any evidence of the influence of the number of choice sets for OHL5.

5. Conclusions[34]

This paper has identified one appealing way of increasing our knowledge on the role of (i) the dimensionality of a stated choice experiment, (ii) the framing of stated choice design profiles relative to an experience profile (a reference base) and (iii) aggregating attributes, in conditioning the processing of the information associated with specific numbers of attributes across a choice set of alternatives in stated choice designs.

The empirical assessment herein, summarized in Table 9, provides evidence on sources of systematic influence on how many attributes are processed relative to the full set offered. This evidence is important in revealing candidate influences on attribute processing. Where we might find evidence of attribute reduction (through exclusion and/or aggregation) as the number of items to process increases (defined as the product of the number of attributes per alternative and number of alternatives), we might reasonably speculate that the selected processing strategy has elements of coping *and* relevancy. This should not necessarily be interpreted as a response to complexity but part of the natural process of decision making.

Table 9 Summary of Candidate Influences on Information Processing

	Natts = 6				Natts = 5			Natts = 4				Natts = 3		
Number Ignored ?	0	1	2	3	0	1	2	0	1	2	3	0	1	2
Design dimensions:														
Number of attribute levels	√	√						√	√		√	√	√	
Number of alternatives	√	√			√	√		√		√		√	√	
Attribute range (narrow)	√	√		√	√	√	√							
Number of choice sets												√	√	
Framing around the Base:														
Reference free flow time vs SC time	√				√	√								
Reference congestion time vs SC time	√				√	√								
Attribute Packaging:														
Adding travel times	√	√			√	√		√			√			
Socio-economics:														
Personal income	√	√				√		√		√		√	√	

In the current study our data does not enable us to separate out the true cognitive burden from spurious cognitive burden, but we do have an increasingly strong belief that strategies centered on coping and relevancy are not dominantly (if at all) a result of SC design complexity, but a result of accumulated experiences in real markets used in dealing appropriately with decisions involving varying amounts of information (in quantity and quality) on offer. Theories of search cost and transactional economics provide support for satisfying type behavior (Simon, 1957) with the number zero on the utility scale being an individual's 'aspiration level'. Thus so long as this level is not reached, an individual will keep experimenting, but once this level is obtained they are satisfied.

[34] This study builds on our ongoing research activity (Hensher 2003, 2003a, in press) as we search for improved ways of understanding how individuals process increasing amounts of information offered to them in stated choice experiments. There is always more to do.

Ongoing research is investigating the role of thinking and memory in developing relevance in SC designs so that the design reflects more meaningfully the context in which individuals are accustomed to making choices. The research of Gilboa and Schmeidler (1995) is particular appealing in its arguments that each choice is evaluated by the sum of the utility levels that resulted from using this choice in past cases, each weighted by the similarity of that past case to the problem at hand (Tversky, 1977). This perspective may call for a revision of the approach in which we design relevancy in the SC experiment that is used to reveal appropriate preferences and choices. We anticipate a pre-phase in which a range of choice experiments are used to reveal the one that reflects most accurately on the preference space within which an individual desires to assess alternatives and arrive at a choice. This will enable us to gain an understanding of 'what matters' in information processing that is truly relevant, in contrast to what is provided that may simply add to cognitive burden and run the risk of being analyzed *as if* it were relevant. Statistical non-significance of a parameter of an attribute should not be seen as necessarily meaning 'not relevant', unless the analyst has been able to quarantine the information that is 'not relevant'. This is a crucial challenge if we are to gain confidence in the behavioral virtues of the SC paradigm.

Unlike the current study where each sampled individual had a fixed SC design (in terms of information quantity) to evaluate, but where the amount varied across the sample (i.e., between-subject variability)[35], ongoing research promotes the need to vary the amount of information within-subject as well to control for differences in response that are confounded by differences in subjects. For example, the same individual should, in a new study, receive a choice task which requires assessment of say, 20 pieces of information (attributes by alternatives), and then a varying amount of information across eight treatments. At each treatment we might ask the individual to indicate how they processed the task in terms of including and excluding particular attributes and the extent to which they aggregated subsets of attributes where this is permissible. At the completion of the choice tasks we might ask the respondent to rank each design in terms of the appropriate amount of information that best represents what they believe is the relevant detail they want to process in order to make a rational choice[36]. This will form the dimensionality of a specific design administered to an individual.

The ongoing research challenge is best stated as follows: What matters is not whether different designs require different information processing strategies, but whether the stated choice design per se contributes to different behavioral responses and associated attribute valuations. Importantly the information processing strategy should be viewed endogenously in future stated choice studies as illustrated in Figure 6.

[35] The variability engendered within and between subjects will provide us with the rich variance required to estimate a discrete choice model capable of separating out the influence of the specific SC tasks dimensionality (in terms of the amount of information to process) on key behavioral outputs such as WTP for specific attributes.
[36] The introduction of a pre-stage of defining the relevant choice set in terms of appropriate amount of information is very appealing. It is not however, the same as adaptive conjoint which uses one design. Rather this focus is on selecting the right design and then preserving d-optimality and other desirable design properties, which are traded away in adaptive designs.

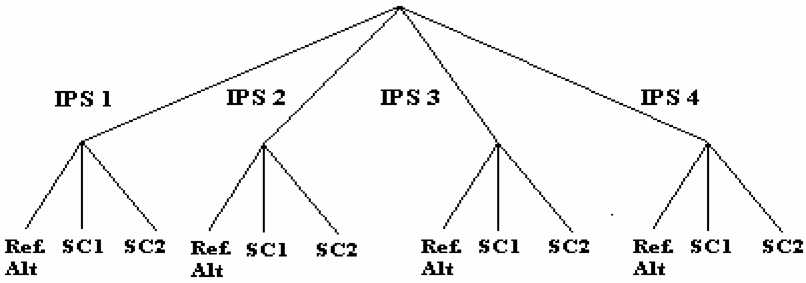

Figure 6 Endogeneity of IPS

Acknowledgement. Discussions over a long period with John Rose, Jordan Louviere, Victor Adamowitz, Joffre Swait and Sean Puckett have added to my understanding of this topic. Specific comments on earlier versions by John Rose and Sean Puckett are appreciated.

6. References

Arentze, T, Borgers, A, Timmermans, H & DelMistro, R (2003) Transport stated choice responses: effects of task complexity, presentation format and literacy, *Transportation Research*, 39e, 229-244.

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

Caussade, S, Ortúzar, Juan de Dios , Rizz, L & Hensher, DA (2003) Assessing the Influence of Design Dimensions on Stated Choice Experiment Estimates, *Design of Designs Report #4,* Institute of Transport Studies, The University of Sydney and Department of Transport Engineering, Pontificia Universidad Católica de Chile.

Dellaert, BGC, Brazell, JD & Louviere, JJ (1999) The effect of attribute variation on consumer choice consistency, *Marketing Letters*, 10, 139-147.

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

Gilboa, I & Schmeidler, D (1995) Case-based decision theory, *Quarterly Journal of Economics*, 110, 605-639.

Gilboa, I, Schmeidler, D & Wakker, P (2002) Utility in case-based decision theory, *Journal of Economic Theory*, 105, 483-502.

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

Hensher, DA (2003) Accounting for Stated Choice Design Dimensionality in willingness to Pay for Travel Time Savings, *Design of Designs Report #2*, Institute of Transport Studies, The University of Sydney, August.

Hensher, DA (2003a) Information processing strategies in stated choice studies: the implications of respondents ignoring specific attributes *Design of Designs Report #3*, Institute of Transport Studies, The University of Sydney, December.

Hensher, DA (in press) Revealing differences in behavioral response due to the dimensionality of stated choice designs: an initial assessment (ARC VTTS Grant 01-03) *Journal of Environmental Resource Economics*, Invited Paper (DoD#1).

Hensher, DA & Goodwin, PB (2004) Implementation of values of time savings: the extended set of considerations in a tollroad context, *Transport Policy*,

Hensher, DA, Rose, J & Greene, WH (2005) *Applied Choice Analysis: A Primer*, Cambridge: Cambridge University Press.

Kahnemann, D & Tversky, A (1979) Prospect theory: an analysis of decisions under risk, *Econometrica*, 47 (2), 263-91.

Klein, NM & Manjit, YS (1989) Context effects on effort and accuracy of choice: an inquiry into adaptive decision making, *Journal of Consumer Research*, 15, 411-421.

Loomes, G & Sugden, R (1987) Some implications of a more general form of regret theory, *Journal of Economic Theory*, 41 (2), 270-87.

Louviere, JJ, Hensher, DA & Swait, JF (2000) *Stated Choice Methods and Analysis*, Cambridge University Press, Cambridge.

Malhotra, NK (1982) Information load and consumer decision making, *Journal of Consumer Research*, 8, 419-430.

Mazzotta, M & Opaluch, J (1995) Decision making when choices are complex: a test of Heiner's hypothesis, *Land Economics*, 71(4), 587-608.

Ohler, T, Li. A, Louviere, J & Swait, J (2000) Attribute range effects in binary response tasks, *Marketing Letters*, 11, 3 249-260.

Rolfe, J, Bennett, J & Blamey, R (2001) Framing effects, in Bennett, J and Blamey, R. (eds), The Choice Modelling Approach to Environmental Valuation, Edward Elgar, Cheltenham, UK, 202-226.

Simon, H (1957) *Models of Man*, John Wiley and Sons, New York.

Starmer, C (2000) Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk, *Journal of Economic Literature*, XXXVIII, 332-382.

Starmer, C & Sugden, R (1993) Testing for juxtaposition and event splitting effects, *Journal of Risk and Uncertainty*, 6, 235-54.

Swait, J & Adamowicz, W (2001) The influence of task complexity on consumer choice: a latent class model of decision strategy switching, *Journal of Consumer Research,* 28, 135-148.

Swait, J & Adamowicz, W (2001a) Choice environment, market complexity, and consumer behavior: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice, *Organizational Behavior and Human Decision Processes*, 49, 1-27.

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

Tversky, A (1977) Features of similarity, *Psychological Review*, LXXXIV, 327-52.

Weber, M, Eisenfuhr, F & Winterfeldt, D (1988) The effects of splitting attributes on weights in multi attribute utility measurement, *Management Science*, 34, 431-45.

White, PJ, Johnson, RD & Louviere, JJ (1998) The effect of attribute range and variance on weighted estimates, unpublished paper, Department of Marketing, The University of Sydney.

Winship, C & Mare, R 1984. Regression models with ordinal variables. *American Sociological Review* 49:512-525.

Appendix A

Table A. Example of a Design Table (One of 16 implemented designs– see Table 2).

Block	Scenarios	Alternative 1			Alternative 2			Alternative 3			Alternative 4		
		Total time	Uncertainty of travel time	Total cost	Total time	Uncertainty of travel time	Total cost	Total time	Uncertainty of travel time	Total cost	Total time	Uncertainty of travel time	Total cost
1	**1**	2	1	2	3	2	1	0	0	3	1	3	0
1	**2**	0	0	1	3	2	2	1	3	0	2	1	0
1	**3**	2	0	0	3	3	1	0	1	2	1	2	1
1	**4**	2	2	0	0	1	3	3	0	1	1	3	2
1	**5**	2	1	1	0	0	2	3	3	0	1	2	0
1	**6**	1	0	2	0	2	0	2	3	3	3	1	1
1	**7**	0	0	1	3	2	0	2	3	0	1	1	0
1	**8**	3	2	3	2	0	3	1	3	1	0	1	2
1	**9**	0	0	3	2	2	2	1	1	1	3	3	2
2	**1**	2	3	1	1	0	3	3	1	0	0	2	2
2	**2**	3	2	3	2	1	3	0	3	3	1	0	2
2	**3**	2	0	1	3	3	0	0	2	0	1	1	2
2	**4**	0	1	3	1	2	1	2	3	0	3	0	2
2	**5**	1	0	3	2	3	1	3	1	2	0	2	3
2	**6**	0	3	1	3	0	0	2	1	3	1	2	2
2	**7**	1	0	3	2	2	2	3	3	0	0	1	1
2	**8**	1	1	3	3	0	3	0	3	2	2	2	1
2	**9**	0	3	0	3	2	1	2	0	2	1	1	3

Appendix B

1. Percentage variations for base range

Table B1.
Values for designs presenting two levels

Label of the attribute	Level 1	Level 2
Free flow time	-20%	20%
Slowed down time	-40%	40%
Stop/start time	-40%	40%
Congestion time	-40%	40%
Total time	-40%	40%
Trip time variability	-40%	40%
Running cost	-20%	20%
Toll cost	-20%	20%
Total cost	-20%	20%

Table B2.
Values for designs presenting three levels

Label of the attribute	Level 1	Level 2	Level 3
Free flow time	-20%	0%	20%
Slowed down time	-40%	0%	40%
Stop/start time	-40%	0%	40%
Congestion time	-40%	0%	40%
Total time	-40%	0%	40%
Trip time variability	-40%	0%	40%
Running cost	-20%	0%	20%
Toll cost	-20%	0%	20%
Total cost	-20%	0%	20%

Table B3.

Values for designs presenting four levels

Label of the attribute	Level 1	Level 2	Level 3	Level 4
Free flow time	-20%	-10%	10%	20%
Slowed down time	-40%	-20%	20%	40%
Stop/start time	-40%	-20%	20%	40%
Congestion time	-40%	-20%	20%	40%
Total time	-40%	-20%	20%	40%
Trip time variability	-40%	-20%	20%	40%
Running cost	-20%	-10%	10%	20%
Toll cost	-20%	-10%	10%	20%
Total cost	-20%	-10%	10%	20%

2. Percentage variations for narrower than base range

Table B4.

Values for designs presenting two levels

Label of the attribute	Level 1	Level 2
Free flow time	-5%	5%
Slowed down time	-20%	20%
Stop/start time	-20%	20%
Congestion time	-20%	20%
Total time	-20%	20%
Trip time variability	-20%	20%
Running cost	-5%	5%
Toll cost	-5%	5%
Total cost	-5%	5%

Table B5.

Values for designs presenting three levels

Label of the attribute	Level 1	Level 2	Level 3
Free flow time	-5%	0%	5%
Slowed down time	-20%	0%	20%
Stop/start time	-20%	0%	20%
Congestion time	-20%	0%	20%
Total time	-20%	0%	20%
Trip time variability	-20%	0%	20%
Running cost	-5%	0%	5%
Toll cost	-5%	0%	5%
Total cost	-5%	0%	5%

Table B6.

Values for designs presenting four levels

Label of the attribute	Level 1	Level 2	Level 3	Level 4
Free flow time	-5%	-3%	3%	5%
Slowed down time	-20%	-3%	3%	20%
Stop/start time	-20%	-3%	3%	20%
Congestion time	-20%	-3%	3%	20%
Total time	-20%	-3%	3%	20%
Trip time variability	-20%	-3%	3%	20%
Running cost	-5%	-3%	3%	5%
Toll cost	-5%	-3%	3%	5%
Total cost	-5%	-3%	3%	5%

3 Percentage values for wider than base range

Table B7.

Values for designs presenting two levels

Label of the attribute	Level 1	Level 2
Free flow time	-20%	40%
Slowed down time	-30%	60%
Stop/start time	-30%	60%
Congestion time	-30%	60%
Total time	-30%	60%
Trip time variability	-30%	60%
Running cost	-20%	40%
Toll cost	-20%	40%
Total cost	-20%	40%

Table B8.

Values for designs presenting three levels

Label of the attribute	Level 1	Level 2	Level 3
Free flow time	-20%	10%	40%
Slowed down time	-30%	15%	60%
Stop/start time	-30%	15%	60%
Congestion time	-30%	15%	60%
Total time	-30%	15%	60%
Trip time variability	-30%	15%	60%
Running cost	-20%	10%	40%
Toll cost	-20%	10%	40%
Total cost	-20%	10%	40%

Table B9.

Values for designs presenting four levels

Label of the attribute	Level 1	Level 2	Level 3	Level 4
Free flow time	-20%	0%	20%	40%
Slowed down time	-30%	0%	30%	60%
Stop/start time	-30%	0%	30%	60%
Congestion time	-30%	0%	30%	60%
Total time	-30%	0%	30%	60%
Trip time variability	-30%	0%	30%	60%
Running cost	-20%	0%	20%	40%
Toll cost	-20%	0%	20%	40%
Total cost	-20%	0%	20%	40%