WORKING PAPER
ITS-WP-04-01

Information Processing Strategies in Stated Choice Studies: The Implications on Willingness to Pay of Respondents Ignoring Specific Attributes

By

David A Hensher

January 2004

Research funded under the Australian Research Council Large Grants Scheme, Grand A00103962

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-04-01

TITLE: Information Processing Strategies in Stated Choice Studies: The Implications on Willingness to Pay of Respondents Ignoring Specific Attributes

ABSTRACT: Individuals processing the information in a stated choice experiment are typically assumed to evaluate each and every attribute offered within and between alternatives and to choose their most preferred alternative. It has always been thought though that some attributes are ignored in this process for many reasons including a coping strategy to handle their perception of the complexity of the choice task. However analysts proceed to estimate discrete choice models as if all attributes have influenced the outcome *to some degree*. The cognitive processes used to evaluate trade-offs are complex with boundaries often placed on the task to assist the respondent. These boundaries can include adding up attributes (eg components of travel time and cost), prioritising attributes and focussing on the primary influences and ignoring specific attributes. In this paper we investigate the implications of bounding the information processing task by attribute elimination through ignoring one or more attributes. Using a sample of car commuters in Sydney we estimate mixed logit models which assume that all attributes are candidate contributors, and models which assume that certain attributes are ignored (based on supplementary information provided by respondents). We derive individual-respondent parameters using a conditional choice specification of mixed logit, and compare the value of travel time savings distribution under alternative information processing regimes. As expected, assuming that all attributes are not ignored and duly processed, leads to biased estimates of parameters which over-estimate willingness to pay (WTP).

KEY WORDS: Stated choice designs, design of designs, information processing, CAPI, behavioural response.

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: January, 2004

1. Introduction

Stated choice (SC) methods are used extensively in many application contexts to reveal the willingness to pay for specific attributes. Within the SC setting, sampled agents typically assess a number of alternatives defined by a set of attributes, each of which is offered as a level drawn from a pre-specified set of levels and range of levels, and they are asked to choose the most preferred alternative (including the choice to not choose any of the offered alternatives). This assessment is repeated a number of times up to the total number of choice sets being offered. The data is then subject to discrete choice modelling using tools such as multinomial logit (MNL), nested logit, mixed logit and latent class MNL.

Stated Choice experiments typically are based on a pre-specified design plan in respect of the number of attributes (including their levels and range), the number of alternatives in a choice set and the number of choice sets to be assessed. While some studies allow for variations in some of these design dimensions, it is common for all sampled agents to be given the exact same number of attributes, alternatives and choice sets. While this is not, per se, a failing of a study, it does raise questions about the influence that the SC design has on the WTP. Without any variation in the dimensionality of the design, it is not possible to assess what influence the design per se has on WTP. Does the design impact in some systematic or non-systematic way on the parameters associated with each attribute? Is the impact stronger in respect of the mean or the variance associated with estimated parameters and the random component of each alternative's utility expression? These important questions have been investigated in Hensher (2003, 2003a) with supporting evidence of design bias.

In Hensher (2003, 2003a) it was assumed, however, that all attributes were deemed relevant in the assessment of the alternatives. To what extent might individuals adopt differing information processing (IP) strategies either to cope with the 'complexity' of an SC experiment and/or because specific attributes are not relevant in their choice? It is reasonable to propose that individuals do have a variety of IP styles, including the *simplifying* strategy of ignoring certain attributes (for whatever reason). Failure to account for such an IP strategy is tantamount to the imposition of the assumption that all designs are comprehensible, all design attributes are relevant (to some degree) and the design has accommodated the relevant amount of complexity necessary to make the choice experiment meaningful. It is important to recognise that simplistic designs may be 'complex' in a perceptual sense, since an individual expects more information which they know is relevant in making such a choice in a real market setting. The development of a series of designs embedded in the one choice experiment, supplemented with questions on how an individual processed the information, enables the researcher to explore sources of systematic influences on choice that if ignored, can lead to biases in key outputs such as willingness to pay.

In this paper we investigate the implications of an individual bounding the information processing task by attribute elimination through ignoring them. In particular we investigate the influence on WTP of an individual stating that they ignored one or more attributes, for whatever reason. We like to think that the attributes were ignored because they are not behaviourally relevant, but we acknowledge that such attributes may be ignored for other reasons (e.g. task simplification).

Using a sample of car commuters in Sydney we estimate mixed logit models which assume that all attributes are candidate contributors, and models which assume that certain attributes are ignored (based on supplementary information provided by respondents). We derive individual-respondent parameters using a conditional choice specification of mixed logit and compare the value of travel time savings distribution under alternative information processing regimes.

2. Information Processing Strategies

Some researchers (eg Heiner 1983) suggest that an increase in choice set complexity will compromise choice consistency, preventing the variation in choice responses being explained by the underlying preference function. This is particularly problematic when the preference function imposes a condition of unlimited human capacity to process information of varying degrees of magnitude and quality in a costless and optimal (minimum effort) manner to arrive at a utility-maximising choice. Heiner argues that increasing choice complexity would widen the gap between an individual's cognitive ability and the cognitive demands of the decision; which would lead to a restriction of the range of decisions considered. While this may satisfy a particular cognitive ability and produce greater predictability in the outcomes, they are not welfare maximising. What we see is an increase in unobserved influences on outcomes (or a relatively higher unexplained error or noise in the random utility function).

In reality individuals adopt a range of bounded rationality conditions as a coping strategy to handle their perception of the complexity of the choice task. The cognitive processes used to evaluate trade-offs are complex with boundaries often placed on the task to assist the respondent. This can include ignoring subsets of attributes, aggregating attributes (where feasible; e.g. components of travel time), imposing thresholds on attribute levels (just noticeable difference), and conditioning one attribute on the level of other attributes. However analysts then proceed to estimate discrete choice models as if all attributes have influenced the outcome *to some degree*.

In previous papers, Hensher (2003, 2003a) has investigated the complexity of the choice task, identified in terms of the number of attributes, the number of choice sets, the number of levels of each attribute and the range of each attribute. Task complexity can be represented by these 'raw' dimensions as well as by a range of representations broadly referred to as information load (a source of cognitive burden). The focus on choice complexity is only interesting when viewed more broadly under what we call the information processing strategy (IPS) of a decision maker. Individual's use a range of IPS's according to their capability to process which is linked to cognitive capability, commitment to effort etc. It is also related to the risk spectrum they wish to operate under, ranging from risk aversion to risk proneness. The greater the risk aversion the smaller the variance in the IPS. The variability in risk is often defined by constructs such as habit formation and variety seeking, both of which suggest mechanisms used to satisfy the individual's commitment of effort and cognitive abilities. If we knew what role these constructs played in behavioural response then we could design an SC experiment tailored to a specific IPS[1].

[1] Such an SC experiment has some similarities to an adaptive choice experiment in which alternative behavioural choice response segments are identified as a way of recognising decision rules such as 'hard-core loyal', 'brand-type', IIA-type and product or service form. This was considered by Kamakura et al

Our challenge becomes the inverse – to have a sufficiently wide ranging set of SC experiments that enable us to reveal the IPS of each decision maker. This setting enables us to investigate the use of attribute elimination (or ignoring an attribute) as a coping strategy or as a genuine process of assessing alternatives and making a choice, and its implications on the value of travel time savings distribution.

3. Revealing Individual Parameters in Mixed Logit

Mixed logit is increasingly used to estimate choice models. There are a number of useful summaries of the method (such as Train (2003) and Hensher and Greene (2003)) and so we will not detail it here. What we do want to do is to discuss the capability of mixed logit to derive observation-specific parameters for random parameters using conditional (on the within sample choice) distributions

In the random utility model of the discrete choice family of models, we assume that a sampled individual ($q = 1,\ldots,Q$) faces a choice among J alternatives in each of T choice situations[2]. The individual is assumed to consider the full set of offered alternatives in choice situation t and to choose the alternative with the highest utility. The (relative) utility associated with each alternative j as evaluated by each individual q in choice situation t is represented in a discrete choice model by a utility expression of the general form in (1).

$$U_{jtq} = \mathrm{b}_q{}'\mathbf{x}_{jtq} + \varepsilon_{jtq} \qquad \textbf{(1)}$$

where $\mathbf{x}_{jtq}$ is a vector of explanatory variables that are observed by the analyst (from any source) and include attributes of the alternatives, socio-economic characteristics of the respondent and descriptors of the decision context and choice task itself (e.g., task complexity in stated choice experiments as defined by number of choice situations, number of alternatives, attribute ranges, data collection method etc) in choice situation t. The components b_q and ε_{jtq} are not observed by the analyst and are treated as stochastic influences.

Within a logit context we impose the condition that ε_{jtq} is independent and identically distributed (IID) extreme value type 1. The IID assumption is restrictive in that its does not allow for the error components of different alternatives to be correlated. We would want to be able to take this into account in some way. One way to do this is to partition the stochastic component additively into two parts. One part is correlated over alternatives and heteroskedastic, and another part is IID over alternatives and individuals as shown in equation (2) (ignoring the t subscript for the present):

$$U_{jq} = \mathrm{b}_q{}'\mathbf{x}_{jq} + (\eta_{jq} + \varepsilon_{jq}) \qquad \textbf{(2)}$$

(1996) as a finite mixture of nested logits (brand and product), latent class (for hard-core) and multinomial logit (IIA) models.

[2] A single choice situation refers to a set of alternatives (or choice set) from which an individual chooses one alternative. They could also rank the alternatives but we focus on first preference choice. An individual who faces a choice situation on more than one occasion (e.g., in a longitudinal panel) or a number of choice sets, one after the other as in stated choice experiments, is described as facing a number of choice situations. Note that the assumption of a fixed choice set size, J, is made purely for convenience at this point; it is inessential.

where η_{jq} is a random term with zero mean whose distribution over individuals and alternatives depends in general on underlying parameters and observed data relating to alternative *j* and individual *q*; and ε_{jq} is a random term with zero mean that is IID over alternatives and does not depend on underlying parameters or data.

The Mixed Logit class of models assumes a general distribution for η_{jq} and an IID extreme value type 1 distribution for ε_{jq} [3]. That is, η_{jq} can take on a number of distributional forms such as normal, lognormal, or triangular. Denote the joint density of $(\eta_{1q}, \eta_{2q},..., \eta_{Jq})$ by $f(\mathrm{h}_q|\mathrm{W})$ where the elements of W are the fixed parameters of the distribution and h_q denotes the vector of *J* random components in the set of utility functions. For a given value of h_q, the conditional probability for choice *j* is logit, since the remaining error term is IID extreme value type 1:

$$L_{jq}(\mathrm{b}_q|\eta_q) = \exp(\mathrm{b}_q'\mathbf{x}_{jq} + \eta_{jq}) / \textstyle\sum_j \exp(\mathrm{b}_q'\mathbf{x}_{jq} + \eta_{jq}). \qquad \textbf{(3)}$$

The unconditional choice probability is this logit probability integrated over all values of h_q weighted by the density of h_q is as shown in equation (4):

$$P_{jq}(\mathrm{b}_q|\mathrm{W}) = \int_{\eta 1q}\int_{\eta 2q} \ldots \int_{\eta Jq} L_{iq}(\mathrm{b}_q|\mathrm{h}_q)\, f(\mathrm{h}_q|\mathrm{W})\, d\eta_{Jq} \ldots d\eta_{2q}\, d\eta_{1q}\,. \qquad \textbf{(4)}$$

Models of this form are called *mixed logit* because the choice probability P_{jq} is a mixture of logits with *f* as the mixing distribution. The probabilities do not exhibit the questionable independence from irrelevant alternatives property (IIA), and different substitution patterns may be obtained by appropriate specification of *f*. This is handled in by a random parameters specification, specifying each element of b_q associated with an attribute of an alternative as having both a mean and a standard deviation. The choice probability in (4) generally cannot be calculated exactly because the integral will not have a closed form. The integral is approximated through simulation (see Train 2003 for more details).

The standard deviation of an element of the b_q parameter vector, which we denote σ_{qk}, accommodates the presence of preference heterogeneity in the sampled population, characterised by an unknown distribution which may take on a user-specified analytical distribution (e.g. normal, triangular).

Of particular interest is the derivation of the conditional individual-specific parameter estimates and the associated values of travel time savings for each individual. As described in Train (2003), we can obtain the conditional estimator for any individual by using Bayes Theorem. The estimator will be (Hensher et al 2003):

[3] The proof in McFadden and Train (2000) that mixed logit can approximate any choice model including any multinomial probit model is an important message. The reverse cannot be said: the multinomial probit model cannot approximate all mixed logit models since the multinomial probit relies critically on normal distributions. If a random term in utility is not normal, then mixed logit can handle it and multinomial probit cannot.

$$\begin{aligned}
\mathrm{E}[\beta_q | \, data_q] &= \int_{\beta_q} \beta_q \, p(\beta_q \,|\, data_q) d\beta_q \\
&= \int_{\beta_q} \beta_q \frac{p(data_q \,|\, \beta_q) \, p(\beta_q)}{p(data_q)} d\beta_q \\
&= \int_{\beta_q} \beta_q \frac{p(data_q \,|\, \beta_q) \, p(\beta_q)}{\int_{\beta_q} p(data_q \,|\, \beta_q) \, p(\beta_q) d\beta_q} d\beta_q \qquad \textbf{(5)} \\
&= \frac{\int_{\beta_q} \beta_q \, p(data_q \,|\, \beta_q) \, p(\beta_q) d\beta_q}{\int_{\beta_q} p(data_q \,|\, \beta_q) \, p(\beta_q) d\beta_q}.
\end{aligned}$$

The conditional density is the contribution of individual q to the likelihood function. The denominator in the conditional mean is the theoretical contribution of individual q to the likelihood function for the observed data, or the conditional choice probability[4]. The numerator of the expectation is a weighted mixture of the values of β_q over the range of β_q where the weighting function is, again, the likelihood function.

Since the integrals cannot be computed analytically, we compute them by simulation. The simulation estimator of the conditional mean for β_q is:

$$\hat{E}_S[\beta_q] = \frac{(1/R)\sum_{r=1}^{R} \beta_{q,r} L(\beta_{q,r} \,|\, data_q)}{(1/R)\sum_{r=1}^{R} L(\beta_{q,r} \,|\, data_q)}. \qquad \textbf{(6)}$$

where the weighting function in each case is the contribution to the likelihood function, computed at the r^{th} draw of $\beta_{q,r}$ in the simulation. The approach in **(6)** can also be used to estimate the conditional variance or standard deviation of β_q by estimating the expected square and subtracting the square of the mean. This estimated conditional variance will be smaller than the average variance obtained simply by computing the sample variance of the estimated conditional means, as the latter is averaged over all the data in the sample while the former is averaged with respect only to the data for individual q.

The distinction betweens conditional and unconditional distributions needs to be explained some more given its importance. A conditional distribution is one in which there are one or more very specific conditions associated with a specific outcome. The conditions of interest are: (i) (***condition A***) a choice probability for an individual that is conditional of the parameter value drawn from the distribution of the error component h (which in the random parameter model produces a parameter value on the distribution) and (ii) (***condition B***) a

[4] Using Bayes Rule, we first define the conditional choice probability as $\mathbf{H}_{jq}(b_q|W) = L_{jq}(b_q)g(b_q|W)/ P_{jq}(b_q|W)$ where $L_{jq}(b_q)$ is now the likelihood of an individual's choice if they had this specific b_q, $g(b_q|W)$ is the distribution in the population of b_qs, and $P_{jq}(W)$ is the choice probability function defined in open-form as (seeTrain (2003):

$$P_{jq}(W) = ?_{b_q} L_{jq}(b_q)g(b_q|W)\, db_q.$$

This shows how one can estimate the person specific choice probabilities as a function of the underlying parameters of the distribution of the random parameters.

parameter estimate associated with an explanatory variable that is influenced by the choice outcome. For mixed logit models, it is possible to have both conditions applying. To make matters slightly confusing, the literature often refers to an unconditional distribution in the context of not taking into account the outcome response (even though the parameter estimates are conditioned on the h (equivalent to condition A).

The most recent advance in the application of mixed logit models is a recognition that the analyst can use the additional information on the observed outcome, as a way of establishing a more behaviourally useful location on the distribution curve for a random parameter to position each individual (in contrast to a random assignment under condition A). The Bayesian literature has always put the view that the inclusion of this additional information (a subjective prior) adds power to the performance of a model (Brownstone 2001). While this position is correct, it is by no means the claim of the Bayesian method per se and indeed is also applicable to classical inference methods that are used in this paper (see Hensher et al 2003 and Huber and Train (2001) for evidence).

4. The Design Plan

Using the context of a car commuter trip, we proposed five design dimensions as the key dimensions of stated choice experiments which are likely to have the greatest contextual influence on choice response and WTP. These dimensions are summarised in Table 1[5].

Table 1 The Dimensions of the Design Plan

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

The SC experiment is 16 Designs embedded in one design each with two versions (ie blocking of 32 rows into sets of 16). Each run of the design determines the specification of a choice experiment that has two versions. For example, the first row might have 15 choice sets of 3 alternatives each presenting 4 attributes at 3 levels. For these specifications an efficient design was created. Six attributes have been selected based on earlier studies (Hensher 2000, 2001). They are: a- free flow time (FFT), b- slowed down time (SDT), c- stop/start time (SST), d- trip time variability (TTV), e- toll cost (TLC), and f- running cost (RC) (based on c/litre, litres/100km). Given that the 'number of attributes' dimension has four levels, we have selected the following combinations of the six attributes, noting that the aggregated attributes are combinations of existing attributes. This is an important point because we did not want the analysis to be confounded by extra attribute dimensions.

[5] 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. Support for the selected dimensions is given in the literature (eg Ohler et al 2000, DeShazo and Fermo 2001, Dellaert et al 1999).

The 3, 4, 5 and 6 attribute levels are:

3: (a+b+c) Total time, d, (e+f) Total costs
4: a,(b+c) Slowed down/Stop Start, d, (e+f)
5: a,b,c,d, (e+f)
6: a,b,c,d,e,f

The master plan gives 16 sub-designs to build as shown in Table 2 representing the attribute profiles in Table 3. Further details are provided in Hensher (2003)

Table 2 The Sub-Designs of the Overall Design

Choice sets 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: The 16 rows represent the set of designs (referred to as Des0,Des1,…..,Des15 in model estimation).

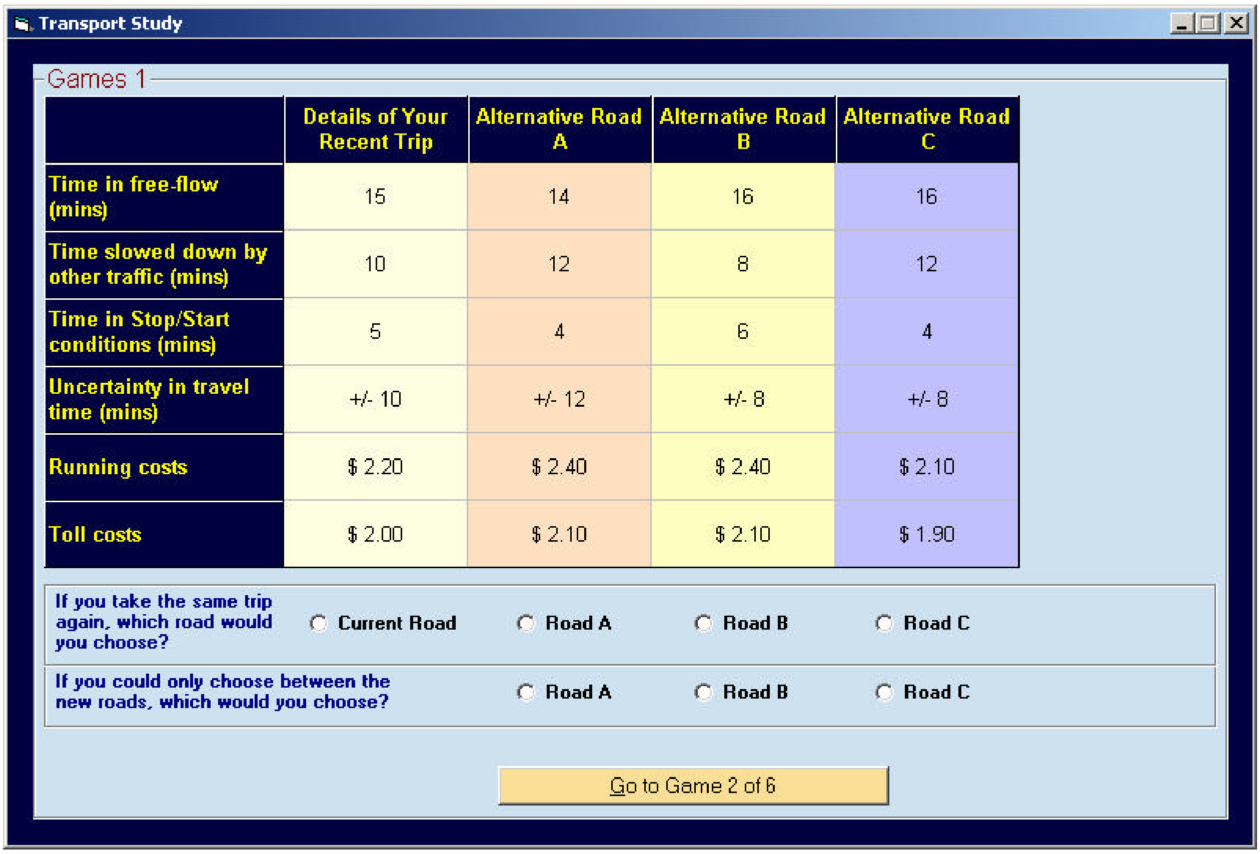
Table 3 Time-Defined attributes and Design Allocation

Design Identifier	Time components (excluding time variability which appears in every design)	Cost components	Number of attributes in design
2	Free flow, slowed down, stop-start	Total cost	5
3	Free flow, slowed down, stop-start	Total cost	5
10	Free flow, slowed down, stop-start	Total cost	5
15	Free flow, slowed down, stop-start	Total cost	5
6	Free flow, slowed down, stop-start	Run cost, toll cost	6
8	Free flow, slowed down, stop-start	Run cost, toll cost	6
9	Free flow, slowed down, stop-start	Run cost, toll cost	6
12	Free flow, slowed down, stop-start	Run cost, toll cost	6
0	Free flow, slowed down-stop-start	Total cost	4
1	Free flow, slowed down-stop-start	Total cost	4
11	Free flow, slowed down-stop-start	Total cost	4
14	Free flow, slowed down-stop-start	Total cost	4
4, 5,7,13	Total time	Total cost	3

The specific design pivots off of the attribute levels associated with a current car-commuting trip. As a generic design we should not expect to find the parameter for each attribute to be different for the full set of alternatives. Thus we can just estimate one parameter per attribute. This reduces the whole design to a set of 8 identical choice sets for each respondent. By doubling the number of choice sets we can allow some 2-way interactions.

The designs are computer-generated. They aim at minimising the correlations between attributes and maximising the amount of information captured by each choice task. Usually, experimental designs are constructed under the assumption of parameters equal to zero. However, to minimise the occurrence of dominant or dominated alternatives, the parameters were assumed to be different from zero. We maximised the determinant of the covariance matrix, which is itself a function of the estimated attribute parameters. Insights from past studies determined their approximate values.

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. The variations used in the choice tasks are given in the Appendix. 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, but the number of attributes and alternatives remain fixed. Variation in the number of attributes and alternatives occurs across commuters.


Transport Study

Games 1

	Details of Your Recent Trip	Alternative Road A	Alternative Road B	Alternative Road C
Time in free-flow (mins)	15	14	16	16
Time slowed down by other traffic (mins)	10	12	8	12
Time in Stop/Start conditions (mins)	5	4	6	4
Uncertainty in travel time (mins)	+/- 10	+/- 12	+/- 8	+/- 8
Running costs	$ 2.20	$ 2.40	$ 2.40	$ 2.10
Toll costs	$ 2.00	$ 2.10	$ 2.10	$ 1.90

If you take the same trip again, which road would you choose?	Current Road	Road A	Road B	Road C
If you could only choose between the new roads, which would you choose?		Road A	Road B	Road C

Go to Game 2 of 6

Figure 1. An example of a stated choice screen

5. Mixed Logit Results

514 face to face CAPI surveys were undertaken in the Sydney metropolitan area in 2002. 502 of the 514 surveys were useable. The 12 rejected surveys were abandoned during the data collection phase due to errors made by interviewers (who entered the data incorrectly onto the CAPI on behalf of the commuter). Full details of the sampling and response rates are given in Hensher (2003).

We estimated mixed logit models in which we (i) did not account for the presence or absence of one of more attributes in information processing and (ii) where we removed an attribute if the individual stated that they ignored it in the assessment of the alternatives. By conditioning the model on whether an attribute is ignored or not, the estimated distribution of parameters becomes the distribution of latent terms that get transformed into the utility function, just like any transformation of a distribution.

The incidence of ignoring one or more attributes is summarised in Table 4. The uncertainty of travel time was ignored by a substantial proportion of the sample (e.g. 36.4% of the individuals facing designs with 5 or 6 attributes) as was total cost (39.35%) for the individuals facing the design with 3 attributes.. In contrast, other mixtures of attributes exhibiting attribute exclusion varied from a high of 11.92% for the pair - slowed down time and stop start time, to a low of 0.49% for mixtures on 3 or 4 time and cost attributes.

Table 4 Profile of Mixtures of Attributes Ignored

Ignored	Proportion
Individuals Facing 5 or 6 attributes (Free flow time, slowed down time, stop start time, uncertainty with run cost, toll cost or total cost):	*Sample =2341*
Free flow time, slowed down time, stop start time	5.28
Slowed down time, stop start time	11.92
Slowed down time	8.05
Free flow time, slowed down time, stop start time, run cost	0.49
Slowed down time, stop start time, run cost	0.98
Free flow time, slowed down time, stop start time, toll cost	0.49
Slowed down time, stop start time, toll cost	0.49
Slowed down time, toll cost	2.52
Free flow time, slowed down time, stop start time, run cost, toll cost	0.49
Uncertainty of time (in presence/absence of other attributes)	36.4
Individuals facing 4 attributes (Free flow time, slowed down time plus stop start time, total cost, uncertainty):	*Sample = 1391*
Free flow time, slowed down time plus stop start time	8.65
Free flow time	3.78
Slowed down time plus stop start time	8.48
Free flow time, slowed down time plus stop start time, total cost	1.62
Free flow time, total cost	1.45
Slowed down time plus stop start time, total cost	1.45
Individuals facing 3 attributes (Total time, total cost, uncertainty):	*Sample = 1035*
Total cost	39.35
Uncertainty	41.52

The mixed logit models in Table 5 used a triangular distribution for the random parameters and constrained the parameters to ensure that the willingness to pay for travel time savings was non-negative. For the triangular distribution, the density function looks like a tent: a peak in the centre and dropping off linearly on both sides of the centre[6].

Two model specifications are reported for each of the models defined on the full sample (NI-1, NI-2) and the sample in which attributes were reported as ignored (I-1, I-2). The model set 1 (NI-1, I-1) specified the travel time component parameters as random whereas the model set 2 (NI-2, I-2) included these same attribute parameters plus the running cost parameter as random. The selection of running cost and not the other costs was designed to illustrate the impact of random parameters on the numerator and denominator of the VTTS formula, to determine if the directional impact of the VTTS in the presence and absence of excluded attributes was impacted on by the selection of only the numerator parameter being random in model set 1. Since we are using individual observation-specific parameters derived from a conditional distribution (condition B) in deriving the VTTS for each individual, we have greater confidence (compared to using the population moments from the unconditional distribution – condition A) in matching the parameters from repeated draws for each of travel time and cost.

Table 5 Mixed Logit Choice Models with alternative information processing conditions (4,593 observations). Time is in minutes, cost is in dollars. (100 Halton draws) Non-ignored refers to all individuals assumed to take the attribute into account.

Attribute	Alternatives	Non-ignored		Ignored	
		NI-1	NI-2	I-1	I-2
Free flow time	2-4, 6-8, 10-12, 14-16, 18-20	-.1583 (-17.9)	-.1586 (-17.9)	-.1588 (-17.4)	-.1591 (-17.3)
Slowed time	3,4, 7,8, 11,12,15,16,19,20	-.1169 (-11.59)	-.1173 (-11.6)	-.1119 (-11.1)	-.1121 (-11.1)
Stop/start time	3,4, 7,8, 11,12,15,16,19,20	-.1499 (-13.82)	-.1503 (-13.8)	-.1482 (-13.7)	-.1484 (-13.6)
Uncertainty time	All	-.0212 (-6.35)	-.0213 (-6.4)	-.0247 (-6.7)	-.0248 (-6.6)
Slowed/stop/start time	2, 6,10,14,18	-.1690 (-17.58)	-.1691 (-17.6)	-.1399 (-15.3)	-.1399 (-15.3)
Total time	1,5,9,13,17	-.1899 (-18.3)	-.1899 (-18.3)	-.1786 (-17.6)	-.1787 (-17.6)
Cost attributes:					
Running cost	4,8,12,16,20	-.8489 (-7.7)	-.8739 (-7.4)	-1.0142 (-7.27)	-1.0468 (-7.0)
Toll cost	4,8,12,16,20	-1.6952 (-24.1)	-1.7043 (-24.1)	-2.3311 (-24.26)	-2.345 (-24.2)
Total cost	1-3,5-7,9-11,13-15,17-19	-1.0661 (-17.8)	-1.067 (-17.8)	-1.4818 (-18.68)	-1.4824 (-18.7)
Standard deviations of random parameters:					
Free flow time	2-4, 6-8, 10-12, 14-16, 18-20	-.1583 (-17.9)	-.1586 (-17.9)	-.1588 (-17.4)	-.1591 (-17.3)
Slowed time	3,4, 7,8, 11,12,15,16,19,20	-.1169 (-11.59)	-.1173 (-11.6)	-.1119 (-11.1)	-.1121 (-11.1)
Stop/start time	3,4, 7,8, 11,12,15,16,19,20	-.1499 (-13.82)	-.1503 (-13.8)	-.1482 (-13.7)	-.1484 (-13.6)
Slowed/stop/start time	2, 6,10,14,18	-.1690 (-17.58)	-.1691 (-17.6)	-.1399 (-15.3)	-.1399 (-15.3)
Total time	1,5,9,13,17	-.1899 (-18.3)	-.1899 (-18.3)	-.1786 (-17.6)	-.1787 (-17.6)
Running cost	4,8,12,16,20		-.8739 (-7.4)		-1.0468 (-7.0)
Pseudo-R^2		0.6485	0.6504	0.6519	0.6519
Log-Likelihood		-4832.78	-4807.6	-4786.49	-4786.2

Notes: All random parameters have a triangular distribution in which the mean = standard deviation. (see Hensher and Greene 2003 for a justification).

[6] Let c be the centre and s the spread. The density starts at c-s, rises linearly to c, and then drops linearly to c+s. It is zero below c-s and above c+s. The mean and mode are c. The standard deviation is the spread divided by $\sqrt{6}$; hence the spread is the standard deviation times $\sqrt{6}$. The height of the tent at c is 1/s (such that each side of the tent has area $s\times(1/s)\times(1/2)=1/2$, and both sides have area 1/2+1/2=1, as required for a density). The slope is $1/s^2$. See Evans et al (1993) for formal proofs.

The overall goodness of fit of all models in impressive, as is often the experience with stated choice studies. All parameters are statistically significant and of the expected sign. The values of travel time savings (VTTS) associated with the conditional distribution are reported in Table 6 for all four models.

As expected, assuming that all attributes are not ignored and duly processed leads to biased estimates of parameters which may under-or over-estimate willingness to pay (WTP) depending on how the bias affects the numerator or denominator of the WTP estimate. In our example where we treat (i) the travel time parameter as random and the cost parameter as fixed and (ii) both time and running cost parameters as random, failure to recognise and account for the information processing strategy of the respondent leads to a significant over-estimate of the mean value of travel time savings, on average of the order of 6-44% depending on the specific attribute. It also impacts on the distribution of VTTS, reducing the variance as well as the minimum and maximum values.

Table 6 Values of travel time savings inclusive and exclusive of individuals who ignored specific attributes

(i) time = random parameter, cost = fixed parameter

Attribute	Non-ignored (NI-1)		Ignored (I-1)	
	Mean (std dev)	Min - Max	Mean (std dev)	Min - Max
Free flow time	11.18 (0.65)	6.02 – 14.90	9.39 (0.49)	5.86 – 12.06
Slowed time	8.26 (0.28)	5.53 – 9.97	6.62 (0.22)	4.54 – 8.44
Stop/start time	10.60 (0.42)	5.33 – 13.20	8.77 (0.33)	4.46 – 11.19
Slowed/stop/start time	11.94 (0.64)	5.79 – 15.29	8.28 (0.34)	4.68 – 10.36
Total time	11.23 (0.69)	3.38 – 15.55	10.57 (0.58)	3.29 – 12.69

(ii) time and running cost = random parameters

Attribute	Non-ignored (NI-2)		Ignored (I-2)	
	Mean (std dev)	Min - Max	Mean (std dev)	Min - Max
Free flow time	10.89 (0.71)	5.85 - 17.0	9.12 (0.56)	5.32 – 15.68
Slowed time	8.06 (0.35)	5.16 – 11.4	6.43 (0.28)	3.78 – 9.75
Stop/start time	10.32 (0.52)	5.82 – 15.6	8.51 (0.42)	4.84 – 15.25
Slowed/stop/start time	11.62 (0.31)	5.29 – 16.05	8.03 (0.41)	4.38 – 11.63

(iii) ratio of non-ignored to ignored mean VTTS

Attribute	Ratio NI/I
Free flow time	1.19
Slowed time	1.25
Stop/start time	1.21
Slowed/stop/start time	1.44
Total time (model set (i) only)	1.06

Note: for all times except total time, running cost is the cost parameter; for total time the cost parameter is total time.

6. Conclusions

The evidence in this paper that a recognition of varying information processing strategies in respect to how specific attributes are processed, in terms of exclusion and inclusion, is compelling. Although we cannot suggest whether the exclusion of an attribute is due to some underlying behavioural rationale for the attribute's role, or simply a coping strategy in processing the amount of information presented in the stated choice experiment, we have empirical evidence which supports the position that imposing a condition of unlimited human capacity to process information of varying degrees of magnitude and quality is not a reflection of how individuals actually make choices. It also artificially produces a willingness to pay distribution for a specific attribute (in our case valuation of travel time savings) that is only true when we assume that all presented attributes matter and individuals are capable of processing the information content of all attributes, as well as wishing to process it.

Indeed in real markets the choice process is simplified for both behavioural and process coping reasons, and as such both sources of potential influence are at play, interacting to produce a specific choice outcome and implied trade-off, and hence valuation of attributes influencing the choice outcome. Accounting for the inclusion vs exclusion of an attribute in an individual's decision calculus does appear to impact significantly on the behavioural outputs of a discrete choice model; in our example the behavioural value of travel time savings distribution and its associated moments are greatly influenced by the assumption made on how attributes are processed.

7. References

Brownstone, D. (2001) Discrete choice modelling for transportation, in Hensher, D.A. (ed) in Hensher, D.A. (ed.) *Travel Behaviour Research: The Leading Edge*, Pergamon Press, Oxford, 97-124.

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

DeShazo, J.R. and G. Fermo (2001) Designing choice sets for stated preference methods: the effects of complexity on choice consistency, *Journal of Environmental Economics and Management*.

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

Hensher, D.A. (2003) Revealing differences in willingness to pay due to the dimensionality of stated choice designs: an initial assessment, *Design of Designs Report #1*, Institute of Transport Studies, The University of Sydney, March.

Hensher, D.A. (2003a) 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, D.A. and Greene, W.H. (2003) Mixed logit models: state of practice, *Transportation*, 30(2), 133-176.

Hensher, D.A., Greene, W. and Rose, J. (2003) Using Classical Inference Methods to reveal individual-specific parameter estimates to avoid the potential complexities of WTP derived from population moments, Institute of Transport Studies, The University of Sydney, September.

Huber, J. and Train, K. (2001) On the similarity of classical and Bayesian estimates of individual mean partworths, *Marketing Letters*, 12, August, 259-270.

Kamakura, W.A., Kim, B-D. and Lee, J. (1996) Modelling preference and structural heterogeneity in consumer choice, *Marketing Science*, 15 (2), 152-172.

McFadden, D. and K. Train (2000) Mixed MNL models for discrete response, *Journal of Applied Econometrics*, 15, 447-470.

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

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

Appendix

1. *Percentage variations for base range*

Table A1.
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 A2.
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 A3.
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 A4.
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 A5.
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 A6.
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 A7.
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 A8.
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 A9.
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%