I T L S

WORKING PAPER
ITLS-WP-07-04

Joint estimation of process and outcome in choice experiments involving attribute framing

By

David A Hensher

March 2007

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in
Transport and Logistics Management

The University of Sydney

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

NUMBER: Working Paper ITLS-WP-07-04

TITLE: Joint Estimation of Process and Outcome in Choice Experiments involving Attribute Framing

ABSTRACT: There is a growing interest and recognition that the study of discrete choice outcomes should take into account the process rules that are used to establish eligibility of each attribute. This applies to both revealed preference and stated choice data but is especially relevant in the context of choice experiments where the analyst traditionally assumes the relevancy of all attributes imposed on the respondent through a series of choice sets. This paper proposes a joint process-outcome model in which the choices made are conditioned on the rules adopted by each respondent in assessing the attributes packaged in the definition of each alternative. We set out a joint model for four attribute processing rules and three alternatives (including a reference alternative), and estimate two sets of panel-based mixed logit models – one set in which we ignore the attribute processing rules and one set in which we explicitly account for the rules. We integrate the inclusion/exclusion rules and 'code' the outcomes of various prospects (i.e., alternatives) as either gains or losses relative to a reference point. Using data from a commuter car trip study of unlabelled packages of times and cost attributes (including a toll), we identify willingness to pay distributions for travel time savings under the various process rules. The main finding is that failing to account for the process rules tends to result in statistically higher mean estimates of values of travel time savings.

KEY WORDS: ***Process, outcome, attribute processing, choice experiments, reference points, willingness to pay, prospects.***

AUTHOR: David A Hensher

CONTACT: Institute of Transport and Logistics Studies **(C37)**
An Australian Key Centre
The University of Sydney NSW 2006 Australia

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

DATE: March 2007

Acknowledgement. Discussions with John Rose and Bill Greene have been invaluable in the preparation of this paper.

1. Introduction

The fast pace at which discrete choice models have been developing since open-form simulation methods became widely available may have left some of the fundamental behavioural building blocks somewhat lagging behind. One area of growing interest is the manner in which we assume individual's process information that is subsequently included in a choice model. Despite a growing number of transportation studies focusing on these issues (see for example Cantillo *et al.* 2006, Hensher 1983, Swait 2001), the entire domain of every attribute is treated as relevant to some degree and included in the utility expressions for every individual. While acknowledging the extensive study of nonlinearity in attribute specification which permits varying marginal (dis)utility over an attribute's range, including account for asymmetric preferences under conditions of gain and loss, this is not the same as establishing ex ante the extent to which a specific attribute might be totally excluded from consideration for all manner of reasons, including the impost of the design of a choice experiment when stated choice data is being used.

The impetus to focus on process rules, treated endogenously, that individuals adopt in assessing a choice experiment and making a choice, is consistent with the contribution that prospect theory has made to understanding behavioural response since Kahneman and Tversky's 1979 pioneering research. Most psychological theories of choice assume a dual-phase model of the decision-making process (Houston *et al.*1989, Kahneman and Tversky, 1979, Thaler, 1999). The first phase relates to the editing of the problem. The second phase relates to the evaluation of the edited problem. The main function of the editing operations is "to organize and reformulate the options so as to simplify subsequent evaluation and choice" (Kahneman and Tversky, 1979, p. 274). The main function of the evaluation operations is to select the preferred alternative. Similarly, in other behavioural paradigms such as the 'Cancellation and Focus Model of Choice' (Houston et al., 1989; Houston and Sherman, 1995, Bonini *et al.* 2004), it is assumed that people cancel features shared by the alternatives (within bounds that allow for just noticeable difference), and focus evaluation on the remaining attributes.

Framing is a critical activity in the construction of behavioural reality that captures many of the elements of stage one editing, including cancellation and focus. Framing helps shape the perspectives through which individuals see the world, focusing attention on key elements within, involving processes of *inclusion* and *exclusion* as well as *emphasis,* and hence operates by biasing the cognitive processes of information by individuals (Hallahan 1999). We are interested in two dimensions of framing – attributes and choice[1]. Attribute framing entails accentuation of attributes of alternatives, ignoring other attributes and hence biasing information processing in terms of focal attributes. Framing of choices entails the posing of alternative outcomes or decisions in either negative (loss) or positive (gain) terms, biasing choices in situations involving uncertainty. Prospect theory suggests that individuals will take greater risks to avoid losses than to obtain gains, and hence the marginal utility of a contrast is asymmetric and nonlinear (see Hess *et al.* 2006). This suggests that circumstances in SC studies where the alternatives on offer that are contrasts to the experienced alternative

[1] Hallahan (1999) presents seven dimensions, of which attributes and choice are only two. The others are situations, actions, issues, responsibility and news.

(e.g., a recent trip) that offer less attractive attribute levels such as travel times, are more likely to result in higher willingness to pay compared to relatively more attractive attribute scenarios.

The role of reference points is central to this formulation of prospects. The framing of alternatives (Tversky and Kahneman 1981) is a major driver of choices. Individuals' 'code' the outcomes of various prospects (i.e., alternatives) as either gains or losses relative to some reference point. This reference point can be an alternative that has been experienced in recent or past times in some accumulative way (as suggested by case-based decision theory – see Gilboa and Schmeidler 1995, Prelec 1998), which is the essence of a pivot-design in stated choice experiments.

The establishment of attribute inclusion/exclusion in making choices in a stated choice (SC) context is often associated with design dimensionality and the so-called *complexity* of the SC experiment (Hensher in press). It is typically implied that designs with more items to evaluate are more complex than those with less items[2] (e.g., Arentze *et al*., 2003, Swait and Adamowicz 2001a, 2001b), impose cognitive burden, and are consequently less reliable, in a behavioural sense, in revealing preference information. This is potentially misleading, since it suggests that complexity is an artefact of the *quantity* of information, in contrast to the *relevance* of information. We need a way of identifying what information (i.e., attributes) is actually processed in arriving at a choice outcome and which is ignored. We recognise however that the process is inherently stochastic from an analyst's perspective, since we will never be able, with total certainty, to rely on a set of exogenous data items to elicit how an attribute is processed by each individual. This necessitates treating the processing of attributes as endogenous with the choice outcome so that unobserved influences of processing can also be accommodated, at least randomly.

This paper proposes a joint process-outcome framing model in which the choices made are conditioned on rules adopted by each respondent in assessing the attributes packaged in the definition of each alternative. We set out a joint model for four attribute processing rules and three alternatives (including a reference alternative), and estimate six mixed logit models – three in which we ignore the attribute processing rules and three in which we explicitly account for the rules. Within each set of three models we identify the influence of asymmetric and non-asymmetric thresholds in gains and losses relative to a reference alternative. The mixed logit model has a flexible structure that can account for between-alternative error structure including correlated choice sets, unobserved preference heterogeneity, and reference dependency. Using data from a non-commuter car trip study of unlabelled packages of times and cost attributes (including a toll), we identify willingness to pay distributions for travel time savings under the various process rules.

[2] 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.

2. The Mixed Logit Process and Outcome Framework

The underlying preference model assumes that individuals evaluate alternatives on offer through an anchoring strategy in which they consider the potential costs and benefits of each alternative relative to the known perceived costs and benefits of experienced alternatives. These alternatives are assumed to be described by a set of attributes explicitly defined by the analyst (the observed attributes) and a set of attributes that matter to the decision maker but which are not explicitly defined by the analyst (the unobserved attributes) in their construction of the alternatives available. Within a population, the observed and unobserved attributes typically have different influences on the choice made and hence preference heterogeneity is all pervasive. Some of this heterogeneity can be related to each observed attribute (through random parameterisation); however there is a real possibility that the attribute package defining each alternative may not be an adequate representation of the preference heterogeneity; resulting in some 'residual' heterogeneity captured by each alternative (through error components) for all manner of underlying reason.

For any one individual and one alternative on offer, the choice model of interest is defined by the joint choice of processing and choosing, which may be structured under the dual-phase model of the decision-making process as the product of the (marginal) choice of attribute processing rule and the outcome choice conditional on the processing rule (Figure 3). Framing is introduced within the joint structure by defining the levels of each attribute offered through each non-experienced alternative (in terms of packaged attribute levels), relative to the reference (experienced) alternative's levels. Reference framing can be defined symmetrically or asymmetrically in terms of the directional magnitude of the difference to signal differential marginal (dis)utility of gains and losses.

The (relative) utility of alternative *j* for individual *i* can be written, assuming linear in parameters, as:

$$U_{ri,apr} = (\alpha_{ri} + \boldsymbol{\beta}_{\mathbf{ri}}\mathbf{X}_{\mathbf{ri}} + \varepsilon_{ri})_{apr}$$

$$U_{ji,apr} = (\alpha_{j} + \beta_{ji}S_j + \boldsymbol{\beta}_{\mathbf{ri}}\mathbf{X}_{\mathbf{ri}} + \boldsymbol{\beta}_{\mathbf{jig}}\mathbf{(X_{ji}\text{-}X_{ri})_{gain}} + \boldsymbol{\beta}_{\mathbf{jil}}\mathbf{(X_{ji}\text{-}X_{ri})_{loss}} + \varepsilon_{ji})_{apr} \tag{1}$$

where $\alpha_{ri,apr}$ is an alternative-specific constant for the reference alternative *r* and individual *i* associated with a specific attribute processing rule *(apr)*; $\mathbf{X_{ri}}$ is a vector of attributes associated with alternative *r* (or *j*) for individual *i*; $\boldsymbol{\beta}_{\mathbf{ri}}$ is a vector of parameters; S_j is a dummy variable (1,0) that tests for sequencing order in the displayed choice set of each non-experienced alternative $\mathbf{(X_{ji}\text{-}X_{ri})_{gain}}$ is a vector of referenced deviations of non-experienced attribute levels $\mathbf{X_{ji}}$, relative to experience levels $\mathbf{X_{ri}}$ for situations where $\mathbf{X_{ji}<X_{ri}}$; $\mathbf{(X_{ji}\text{-}X_{ri})_{loss}}$ is the equivalent reference deviation when $\mathbf{X_{ji}>X_{ri}}$[3] ; and ε_{ji} is a random component that captures through a series of assumptions (see below) the unobserved sources of preference heterogeneity that can be ascribed to attributes and alternatives. Within the mixed logit framework, random taste

[3] The model can also account for the situation where $X_{ji}=X_{ri}$.

heterogeneity can be aligned to attributes through random parameters and to alternatives through error components.

Utility expressions are defined for each choice set where a choice set has alternatives representing a specific attribute processing rule (APR). For example, one APR might have all attributes included in making a choice; another APR might exclude a particular attribute. The likelihood of the universal set of APR's and choices sets from which an alternative is chosen is computed jointly across all choice sets. The full mixed logit model with all components in choice setting t is given in (2).

$$\text{Prob}(y_{it} = j) = \frac{\exp\left[\alpha_{ji} + \boldsymbol{\beta}_i'\mathbf{x}_{jit} + \Sigma_{m=1}^{M} d_{jm}\theta_m E_{im}\right]}{\sum_{q=1}^{J_i}\exp\left[\alpha_{qi} + \boldsymbol{\beta}_i'\mathbf{x}_{qit} + \Sigma_{m=1}^{M} d_{qm}\theta_m E_{im}\right]} \quad (2)$$

$(\alpha_{ji}, \boldsymbol{\beta}_i) = (\alpha_j, \boldsymbol{\beta}) + \boldsymbol{\Gamma\Omega}_i\mathbf{v}_i$ are random alternative-specific constants and taste parameters; $\boldsymbol{\Omega}_i$ = diag(ω_{i1}, ω_{i2}, ...) or $\boldsymbol{\Omega}_i$ = diag(σ_1,...,σ_k); and $\boldsymbol{\beta}$,α_{ji} are constant terms in the distributions of the random taste parameters. Uncorrelated parameters with homogeneous means and variances are defined by $\beta_{ik} = \beta_k + \sigma_k v_{ik}$ when $\boldsymbol{\Gamma} = \mathbf{I}$, $\boldsymbol{\Omega}_i$ = diag(σ_1,...,σ_k), $\mathbf{x}_{jit}$ are observed choice attributes and individual characteristics, and $\mathbf{v}_i$ is random unobserved taste variation, with mean vector **0** and covariance matrix **I.** This model accommodates correlated parameters with homogeneous means through defining $\beta_{ik} = \beta_k + \Sigma_{s=1}^{k}\Gamma_{ks} v_{is}$ when $\boldsymbol{\Gamma} \neq \mathbf{I}$, and $\boldsymbol{\Omega}_i$ = diag(σ_1,...,σ_k), with Γ defined as a lower triangular matrix with ones on the diagonal that allows correlation across random parameters when $\boldsymbol{\Gamma} \neq \mathbf{I}$. An additional layer of individual heterogeneity can be added to the model in the form of the error components. The individual specific underlying random error components are introduced through the term E_{im}, m = 1,...,M, E_{im} ~ N[0,1], given d_{jm}= 1 if E_{im} appears in utility for alternative j and 0 otherwise, and θ_m is a dispersion factor for error component m.

The probabilities defined above are conditioned on the random terms, $\mathbf{v}_i$ and the error components, $\mathbf{E}_i$. The unconditional probabilities are obtained by integrating v_{ik} and E_{im} out of the conditional probabilities: $P_j = E_{\mathbf{v},\mathbf{E}}[P(j|\mathbf{v}_i,\mathbf{E}_i)]$. This is a multiple integral which does not exist in closed form. The integral is approximated by sampling *nrep* draws from the assumed populations and averaging. See Bhat (2003), Revelt and Train (1998), Train (2003) and Brownstone *et al.* (2000) for discussion. Parameters are estimated by maximizing the simulated log likelihood given in (3).

$$\log L_s = \sum_{i=1}^{N} \log\frac{1}{R}\sum_{r=1}^{R}\prod_{t=1}^{T_i}\frac{\exp\left[\alpha_{ji} + \boldsymbol{\beta}_{ir}'\mathbf{x}_{jit} + \Sigma_{m=1}^{M} d_{jm}\theta_m E_{im,r}\right]}{\sum_{q=1}^{J_i}\exp\left[\alpha_{qi} + \boldsymbol{\beta}_{ir}'\mathbf{x}_{qit} + \Sigma_{m=1}^{M} d_{qm}\theta_m E_{im,r}\right]} \quad (3)$$

with respect to ($\boldsymbol{\beta}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\Omega}$, $\boldsymbol{\theta}$), where R = the number of replications, $\boldsymbol{\beta}_{ir} = \boldsymbol{\beta} + \boldsymbol{\Gamma\Omega}_i\mathbf{v}_{ir}$ is the rth draw on $\boldsymbol{\beta}_i$, $\mathbf{v}_{ir}$ is the rth multivariate draw for individual i, and $E_{im,r}$ is the rth univariate normal draw on the underlying effect for individual i. The multivariate draw, v_{ir} is actually K independent draws. Heteroscedasticity is induced first by multiplying by Ω_i, then the correlation is induced by multiplying $\Omega_i v_{ir}$ by Γ.

The alternative-specific constants in (3) are linked to the EV1 type distribution for the random terms, after accounting for unobserved heterogeneity induced via distributions imposed on the observed attributes, and the unobserved heterogeneity that is alternative-specific and accounted for by the error components. The error components account for correlated observations across choice sets as well as unobserved (to the analyst) differences across decision-makers in the intrinsic preference for a choice alternative (or preference heterogeneity). The parameter associated with each error component is $\beta*\sigma$, neither of which appears elsewhere in the model. We induce meaning by treating this parameter pair as θ which identifies the variance of the alternative-specific heterogeneity. What we are measuring is variation around the mean[4].

3. Empirical Application

The data used to estimate the models is drawn from a study undertaken in Sydney in 2004, in the context of car driving non-commuters making choices from a range of level of service packages defined in terms of travel times and costs, including a toll where applicable. The sample of 189 effective interviews, each responding to 16 choice sets, resulted in 3,024 observations for model estimation[5].

To ensure that we captured a large number of travel circumstances and potential attribute processing rules, that will enable us to see how individuals trade-off different levels of travel times with various levels of tolls, we sampled individuals who had recently undertaken trips of various travel times, in locations where tollroads currently exist.[6] To ensure some variety in trip length, three segments were investigated: no more than 30 minutes, 31 to 60 minutes, and more than 61 minutes (capped at two hours).

A telephone call was used to establish eligible participants from households stratified geographically, and a time and location agreed for a face-to-face computer aided personal interview (CAPI). A stated choice (SC) experiment offers the opportunity to establish the preferences of travellers for existing and new route offerings under varying packages of trip attributes. The statistical state of the art of designing SC experiments has moved away from orthogonal designs to statistically efficient designs (see Rose *et al.* 2005, Johnson *et al.* 2006, Sandor and Wedel 2002); and the behavioural state of the art has moved to promoting designs that are pivotted around the knowledge base of travellers, in recognition of supporting theories in behavioural and cognitive psychology and economics such as prospect theory, case-based decision theory and minimum-regret theory. Starmer (2000, p 353) makes a very strong plea in support of the use of reference points (i.e., a current trip):

---

4 The idea that beta is the coefficient on the unmeasured heterogeneity might be strictly true, but the concept does not work in other models that have error components in them, so we should not try to impose it here. For example, in the linear model, we have an unmeasured variable epsilon, and we write the model y = a +x'b + sigma*epsilon where, strictly speaking, epsilon is the unmeasured heterogeneity and sigma is the coefficient. But, sigma is the standard deviation of the unmeasured heterogeneity, not the "coefficient" on the unmeasured heterogeneity.

5 This is not the entire data set. It is 85 percent of the sample, selected as the sub-sample that chose one of the four attribute processing rules included herein. There were a number of other rules spread across the balance of the sample (see Hensher 2006).

6 Sydney has a growing number of operating tollroads; hence drivers have had a lot of exposure to paying tolls. Indeed, Sydney has the greatest amount of urban kilometres under tolls than any other metropolitan area with the possible exception of Santiago.

> *"While some economists 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"*

A statistically efficient design is a design that minimizes the elements of the asymptotic (co)variance matrix, Ω, with the aim of producing greater *reliability* in the parameter estimates given a fixed number of choice observations. To compare the statistical efficiency of SC experimental designs, a number of alternative approaches have been proposed within the literature (see e.g., Bunch *et al.* 1996). The most commonly used measure is D-error.

$$\text{D-error} = \left(\det \Omega\right)^{1/k} = -\frac{1}{N}\left(\det\left(\frac{\partial LL(\beta)^2}{\partial\beta\partial\beta'}\right)\right)^{-1/k}. \quad (4)$$

where k represents the number of parameters for the design, $LL(\beta)$ the log-likelihood function of the discrete choice model under consideration, N the sample size, and β the parameters to be estimated from the design. Given that we are generating designs and not estimating parameters for an already existing design, it is necessary to assume a set of priors for the parameter estimates. Given uncertainty as to the actual population parameters, it is typical to draw these priors from Bayesian distributions rather than assume fixed parameter values. Typically normal and uniform Bayesian distributions are used (uniform distributions are used if the direction and magnitude of the parameter estimates are unknown; e.g., Kessel *et al.* 2006).

The $D_{(b)}$-error is calculated by taking the determinant, with both scaled to take into account the number of parameters to be estimated. It involves a series of multiplications and subtractions over all the elements of the matrix (see for example, Kanninen 2002). As such, the determinant (and by implication, the $D_{(b)}$-error measure) summarizes all the elements of the matrix in a single 'global' value. Thus, whilst attempts to minimize the D-error measure, on average, minimize all the elements within the matrix, it is possible that in doing so, some elements (variances and/or covariances) may in fact become larger. Despite this property, the $D_{(b)}$-error measure has become the most common measure of statistical efficiency within the literature.

The two stated choice alternatives are unlabelled routes. The trip attributes associated with each route are summarised in Table 1. These were identified from reviews of the literature and through the effectiveness of previous VTTS studies undertaken by Hensher (2001).

Table 1: Trip Attributes in Stated Choice Design

Routes A and B
Free flow travel time
Slowed down travel time
Trip travel time variability
Running cost
Toll Cost

All attributes of the SC alternatives are based on the values of the current trip. Variability in travel time for the current alternative was calculated as the difference between the longest and shortest trip time provided in non-SC questions. The SC alternative values for this attribute are variations around the total trip time. For all other attributes, the values for the SC alternatives are variations around the values for the current trip. The variations used for each attribute are given in Table 2.

Table 2: Profile of the Attribute range in the SC design

	Free-flow time	Slowed down time	Variability	Running costs	Toll costs
Level 1	- 50%	- 50%	+ 5%	- 50%	- 100%
Level 2	- 20%	- 20%	+ 10%	- 20%	+ 20%
Level 3	+ 10%	+ 10%	+ 15%	+ 10%	+ 40%
Level 4	+ 40%	+ 40%	+ 20%	+ 40%	+ 60%

The experimental design has one version of 16 choice sets (games). The design has no dominance given the assumptions that less of all attributes is better. The distinction between free flow and slowed down time is designed to promote the differences in the quality of travel time between various routes – especially a tolled route and a non-tolled route, and is separate to the influence of total time. Free flow time is interpreted with reference to a trip at 3 am in the morning when there are no delays due to traffic.[7] An example of a stated choice screen is shown as Figure 1 with elicitation questions associated with attribute inclusion and exclusion shown in Figure 2.

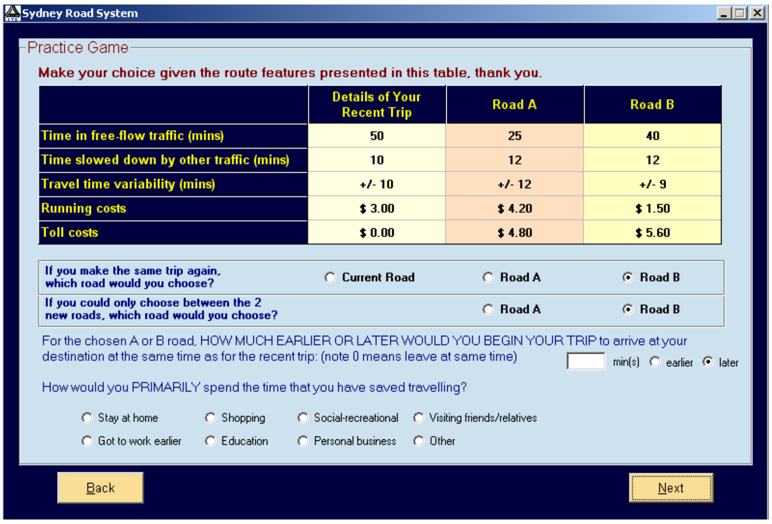

Figure 1: An example of a stated choice screen

[7] This distinction does not imply that there is a specific minute of a trip that is free flow per se but it does tell respondents that there is a certain amount of the total time that is slowed down due to traffic etc and hence a balance is not slowed down (i.e., is free flow like one observes typically at 3am in the morning).

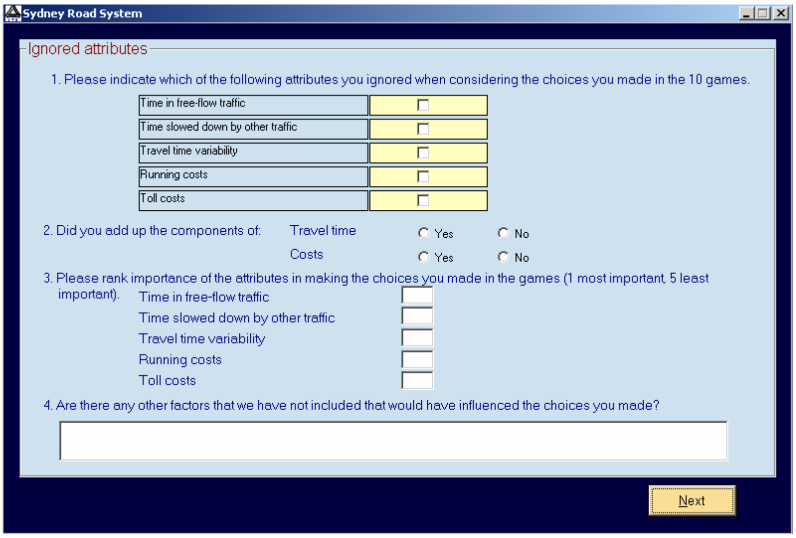

Figure 2: CAPI questions on attribute relevance

4. Empirical Results

The incidence of attribute processing rules is given in Table 3. Well over half (i.e., 62.4 percent) of the sample attended to every attribute and not one respondent attended to none of the attributes. Running cost was the least attended to attribute when one attribute was ignored (i.e., 21.1 percent of the sample); in contrast the toll cost was attended to for 93.7 percent of the sample. Free flow and slowed down were not attended to by 10 percent of the sample. Hence 90 percent of the sample attended to the components of travel time and 72.5 percent attended to the components of cost. It should be noted that these processing rules are assumed to be context-specific and may not apply when an individual is faced with a different attribute dimensionality. For example, individuals who excluded free flow and slowed down time were faced with a higher toll regime than the other respondents, which may have resulted in a dominating focus on the toll payment[8].

Table 3: Incidence of Mixtures of Attributes Processed

Attribute Processing Profile	*Attribute Processing Rule*	*Sample no. of observations=3024*
All attributes attended to	APR1	1888
Attributes not attended to:		
Running cost	APR2	640
Running and toll cost	APR3	192
Free flow and slowed down time	APR4	304

8 In future studies we recommend that questions are asked to elicit the bounds, or caps, that individuals place on specific attributes, if the attribute is to be a candidate for assessment. Swait (2001) asked such questions in the context of purchasing specific vehicle types.

Six mixed logit models have been estimated to investigate the roles of attribute processing rules, asymmetric gains and losses in marginal (dis)utility and non-asymmetry in marginal (dis)utility around the reference alternative. The base model ignores the heterogeneity in attribute processing rules (assuming a single APR that all attributes are relevant), and does not account for the possibility of bias due to reference dependency. Given an unlabelled set of alternatives, the parameters associated with each of the travel time and cost attributes are specified as generic. These findings are summarized in Tables 4 and 5[9]. Each model has an associated set of values of travel time savings, shown in Table 6. All random parameters have been specified as constrained triangular distributions[10].

The base model has five random parameters, two non-random parameters and three error components. All parameter estimates are statistically significant. The reference alternative-specific constant has the expected positive sign, indicating an inertia bias towards the experienced alternative after accounting for the levels of all attributes, all assumed to be relevant to some degree. The error components distinguish between each of the three alternatives, with a greater amount of unobserved heterogeneity associated with the reference alternative. For the SC alternatives, we find that the degree of unobserved heterogeneity is greater for the alternative closer to the reference alternative, after controlling for the observed attributes and alternative-specific sequencing through inclusion of an SC-specific dummy variable for the middle alternative[11]. This dummy variable has a positive sign, suggesting some inertia bias in the second alternative relative to the third alternative (possibly due to 'closeness' on the CAPI screen to the reference alternative), after allowing for the reference alternative-specific effect.

[9] Models in which we allowed a fully generic specification for an APR model (with times treated as random parameters and costs with both random and non-random parameter specification) were estimated and are available on request. A similar model is available for the non-APR structure.

[10] The triangular distribution was first used for random coefficients by Train and Revelt (2000) and Train (2001), later incorporated into Train (2003). Hensher and Greene (2003) also used it and it is increasingly being used in empirical studies. 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 $1/s^2$. The mean weighted average elasticities were also statistically equivalent.

[11] The significance of an ASC related to an unlabelled alternative simply implies that after controlling for the effects of the modelled attributes, this alternative has been chosen more or less frequently than the base alternative. It is possible that this might be the case because the alternative is close to the reference alternative, or that culturally, those undertaking the experiment tend to read left to right. Failure to estimate an ASC would in this case correlate the alternative order effect into the other estimated parameters, possibly distorting the model results.

Table 4 Model Results for Joint Process and Choice Outcome Mixed Logit Models without Thresholds, 3024 observations (RP = random parameter, NRP = non-random parameter), 500 Halton draws

Attribute	Ignoring Process Rule	Accounting for Process Rule
Mean of random parameters:	Generic RP time and cost	APR-Specific RP Time and Cost
Free flow time	-.1363 (-24.5)	
Free flow time APR1		-.1117 (-14.2)
Free flow time APR2		-.1779 (-14.3)
Free flow time APR3		-.1649 (-5.03)
Slowed down time	-.1342 (-21.0)	
Slowed down time APR1		-.1475 (-19.9)
Slowed down time APR2		-.1361 (-7.59)
Slowed down time APR3		-.1180 (-2.39)
Running cost	-.5637 (-21.5)	
Running Cost APR1		-.7346 (-15.8)
Running Cost APR4		-.3946 (-4.06)
Toll cost	-1.115 (-22.1)	
Toll Cost APR1		-1.358 (-17.5)
Toll Cost APR2		-.6192 (-5.85)
Toll Cost APR4		-.5356 (-1.95)
Toll-specific constant	1.4361 (8.14)	
Toll-specific constant APR1		2.354 (8.73)
*Std Dev random parameters:**		
Free flow time	-.1363 (-24.5)	
Free flow time APR1		-.1117 (-14.2)
Free flow time APR2		-.1779 (-14.3)
Free flow time APR3		-.1649 (-5.03)
Slowed down time	-.1342 (-21.0)	
Slowed down time APR1		-.1475 (-19.9)
Slowed down time APR2		-.1361 (-7.59)
Slowed down time APR3		-.1180 (-2.39)
Running cost	-.5637 (-21.5)	
Running Cost APR1		-.7346 (-15.8)
Running Cost APR4		-.3946 (-4.06)
Toll cost	-1.115 (-22.1)	
Toll Cost APR1		-1.358 (-17.5)
Toll Cost APR2		-.6192 (-5.85)
Toll Cost APR4		-.5356 (-1.95)
Toll-specific constant	1.4361 (8.14)	
Toll-specific constant APR1		2.354 (8.73)
Non-random parameters:		
Reference alt constant	.3676 (3.51)	
Reference alt constant APR1		-.0300 (-.22)
Reference alt constant APR2		.8301 (2.80)
Reference alt constant APR3		-.2729 (-.32)
Reference alt constant APR4		1.2048 (1.41)
SC alt 2 constant	.1959 (2.37)	
SC constant APR1		.2139 (3.37)
SC constant APR2		.3372 (1.69)
SC constant APR3		-.9571 (-2.90)
SC constant APR4		.2968 (1.15)
Toll-specific constant APR2		-.6602 (-1.47)
Toll-specific constant APR3		-.8838 (-1.55)
Toll-specific constant APR4		-.0299 (-.02)
Error Component (Alt specific heterogeneity):		
Reference Alternative	2.6505 (20.8)	
SC Alternative 1	.4729 (5.52)	
SC Alternative 2	.3049 (3.06)	
Reference alts (1,2,3,4)		1.5336 (11.9)
SC alt APR1		.8094 (5.06)
SC alt APR2		.2932 (1.0)
SC alt APR3		1.908 (2.39)
SC alt APR4		2.5175 (3.4)
Log-likelihood at convergence	-1757.03	-1760.19

Note: * = constrained triangular random parameter.

The base model alternatives are conditioned on the attribute processing rules in order to identify the extent to which specific attribute effects vary according to the processing rule. With four APR's we do indeed find some statistically different impacts. For free flow time, the marginal disutility is highest when running cost is excluded (AP2), and slightly lower when both running and toll cost are excluded (APR3), in contrast to when all attributes are being traded (APR1). Slowed down time, however, declines in marginal disutility as we move from APR1 through to APR3, and toll cost marginal disutility declines as we move from APR1, to APR2 to APR4. A similar finding applies to running cost in moving from APR1 to APR4. This suggests that there are noticeable differences in the marginal disutilities of each attribute under alternative attribute processing rules. In particular a specific attributes marginal disutility is clearly influenced by the trade-offs being assessed between specific attributes, with strong evidence that the marginal disutility of an attribute decreases as the number of attributes to be processed decreases. Another way of stating this is that the assumption that each and every attribute is relevant tends to overstate the marginal disutility of an attribute, in comparison to when the process of inclusion and exclusion is invoked through framing under stage one editing.

Taking a closer look at the reference alternative-specific constant, which is statistically significant under the full relevancy assumption, it looses its statistical significance under all attribute processing rules except where running cost is excluded (APR2). Although the reasoning is not obvious, this may in part be due to the fact that invoking of the APR itself positions each attribute across the alternatives in decision space that enables less reliance on the reference alternative in order to account for inertia bias. That is, once we remove attributes that simply do not matter, individuals pay closer attention to the actual attribute levels across all alternatives, and hence rely less on referencing of the experienced alternative in exercising their choices. This suggests that assuming attribute relevancy for all attributes artificially forces a disproportionate reliance on the base alternative to assist in choosing.

Finally the error component parameter estimates suggest the existence of greater unobserved heterogeneity under APR's when more attributes are excluded, especially the exclusion of the two travel time attributes, implying that a number of other influences may be at play here. This is consistent with a view, reflected in the levels of toll and running cost, that individuals who exclude free flow and slowed down time are focusing on the trip costs that are especially high, and hence the travel time differences are of little consequence in the choice. This inference may signal a concern about the levels of times and costs that should be included in a choice experiment if we are to be able to infer time-cost trade offs and hence values of travel time savings. In the current study, APR3 and APR4, which are 16.4 percent of the sample, do not enable derivation of time-cost trade-offs.

The models in Table 5 extend the base non-APR and APR models in Table 4 to account for symmetric and asymmetric marginal disutility variations associated with attribute levels of the non-reference alternatives, compared to levels associated with the reference alternative. Separate parameters are estimated for increases and decreases in an attribute in relation to the reference alternative, allowing for asymmetrical responses to be captured through the marginal disutilities of gains and losses in each attribute level.

There is clear evidence that asymmetry exists, although the referencing has a mixed statistical significance across attributes and processing rules. In the context of ignoring the heterogeneity in processing rules, free flow travel time is statistically significant under symmetry and asymmetry, with the marginal utility being higher under gains than under losses. This implies that the marginal disutility of free flow time is discounted less when the non-reference alternatives' free flow time is higher than that of the reference alternative, compared to when it is lower. The only other referencing effect that is statistically significant is slowed down time under losses. For running cost and toll cost, we find that referencing has no statistically significant effect on the role of these two attributes in choosing amongst the three alternatives. These findings also apply when the attribute processing rules are included endogenously. This evidence supports the view that ideas promoted in prospect theory of symmetric and asymmetric framing are not always supported for all attributes, but appear to be attribute-specific and independent of whether the inclusion/exclusion process rules are in place.

We find that the inertia parameter associated with the reference alternative is statistically significant under asymmetry referencing but not under symmetry referencing; however the significance when APR heterogeneity is not accounted for masks the empirical evidence that this is due solely to the statistical significance under the attribute processing rule where all attributes are stated as relevant (i.e., APR1). Again, this supports the finding in Table 4 for the base models that when we account for the attribute processing rules, reference inertia is no longer statistically significant under rules that accommodate attribute irrelevance (i.e., exclusion). This has important implications on behavioural outputs such as willingness to pay (see below), given that the reference constant cannot be deemed to be a proxy for the APR.

The error components are statistically significant under asymmetric referencing when attribute processing is treated heterogeneously and otherwise, and also under heterogeneous APR for symmetrical referencing. This suggests that there exist statistically significant sources of unobserved heterogeneity that are alternative-specific, which vary according to the attribute processing rule and whether the alternative is a reference or non-reference alternative. Most notable is the evidence that unobserved heterogeneity is greatest for the non-reference alternatives across all APR's, in contrast to the reference alternatives associated with each APR. This is a very important finding, highlighting the differences in error conditions attached to non-reference alternatives that are processed differently after separating out the special role that the reference alternative plays. The error components presented for the models that ignore the four attribute processing rules, appear to confound the role of APR's and in particular upwardly bias the unobserved heterogeneity associated with the reference alternatives. Again we find further evidence of the role of attribute processing rules in reducing the inertia bias attributed to the reference alternative, but this time it is associated with the variation in unobserved sources of alternative-specific utility for each APR, in contrast to the mean effects bias in the reference alternative-specific constants.

There are many possible candidate reasons for a particular APR. These include having a mix of levels of attributes that makes certain attributes inconsequential in presence of levels of other attributes (e.g., a very high toll relative to perceived gains in travel time), an opposition to tolling (linked to private concessions), and a broad position on the mixture of levels of attributes that would be acceptable in real markets. The latter point relates to bounded rationality and the risk that SC experiments offer attribute level mixes that are outside of the acceptable bounds.

Table 5 Model Results for Joint Process and Choice Outcome Mixed Logit Models with Referencing, 3024 observations (RP = random parameter, NRP = non-random parameter), 500 Halton draws

Attribute	Ignoring Process Rule		Accounting for Process Rule	
Mean of random parameters:	Non-asymmetric referencing	Asymmetric referencing (gains and losses)	Non-asymmetric referencing	Asymmetric referencing (gains and losses)
Free flow time	-.1734 (-20.72)	-.17779 (-18.7)		
Free flow time APR1			-.1384 (-11.5)	-.1579 (-13.4)
Free flow time APR2			-.2188 (-8.8)	-.1857 (-10.5)
Free flow time APR3			-.1918 (-5.1)	-.1845 (-5.5)
Slowed down time	-.1652 (-16.30)	-.1518 (-15.47)		
Slowed down time APR1			-.1756 (-13.3)	-.1692 (-11.9)
Slowed down time APR2			-.1191 (-7.3)	-.1366 (-5.95)
Slowed down time APR3			-.1223 (-1.54)	-.1083 (-1.39)
Running cost	-.7259 (-11.2)	-.7708 (-11.5)		
Running Cost APR1			-.7914 (-9.8)	-.7830 (-9.3)
Running Cost APR4			-.6456 (-4.1)	-.4913 (-3.45)
Toll cost	-1.3288 (-16.6)	-1.2047 (-16.4)		
Toll Cost APR1			-1.3010 (-12.5)	-1.412 (-10.9)
Toll Cost APR2			-.9697 (-4.9)	-.9577 (-4.9)
Toll Cost APR4			-.6111 (-2.9))	-.4534 (-2.8)
Toll-specific constant	2.2178 (7.91)	1.9294 (6.33)		
Toll-specific constant APR1			1.998 (6.0)	2.416 (4.87)
Reference Points:				
Free flow (SC-Ref)	.0279 (5.73)		.0186 (2.8)	
Slowed down time (SC-Ref)	.0100 (1.92)		.0074 (1.3)	
Running cost (SC-Ref)	.0466 (1.32)		.0516 (1.22))	
Toll cost (SC-Ref)	.0196 (.85)		.0028 (.13)	
Free flow (SC-Ref) -Gains		.0326 (4.2)		.0276 (2.78)
Slowed down time (SC-Ref) - Gains		-.0075 (-.76)		-.0121 (-1.16)
Running cost (SC-Ref) - Gains		.0895 (1.63)		-.0005 (-.01)
Toll cost (SC-Ref) - Gains		.0175 (.45)		-.0307 (-.74)
Free flow (SC-Ref) - Losses		.0198 (2.71)		.0163 (2.02)
Slowed down time (SC-Ref) - Losses		.02169 (2.43)		.0171 (1.92)
Running cost (SC-Ref) - Losses		.1119 (1.47)		.0729 (.86)
Toll cost (SC-Ref) - Losses		.0248 (.80)		.0036 (.10)
*Std Dev random parameters:**				
Free flow time	-.1734 (-20.72)			
Free flow time APR1			-.1384 (-11.5)	-.1579 (-13.4)
Free flow time APR2			-.2188 (-8.8)	-.1857 (-10.5)
Free flow time APR3			-.1918 (-5.1)	-.1845 (-5.5)
Slowed down time	-.1652 (-16.30)			
Slowed down time APR1			-.1756 (-13.3)	-.1692 (-11.9)
Slowed down time APR2			-.1191 (-7.3)	-.1366 (-5.95)
Slowed down time APR3			-.1223 (-1.54)	-.1083 (-1.39)
Running cost	-.7259 (-11.2)			
Running Cost APR1			-.7914 (-9.8)	-.7830 (-9.3)
Running Cost APR4			-.6456 (-4.1)	-.4913 (-3.45)
Toll cost	-1.3288 (-16.6)			
Toll Cost APR1			-1.3010 (-12.5)	-1.412 (-10.9)
Toll Cost APR2			-.9697 (-4.9)	-.9577 (-4.9)
Toll Cost APR4			-.6111 (-2.9))	-.4534 (-2.8)
Toll-specific constant	2.2178 (7.91)			
Toll-specific constant APR1			1.998 (6.0)	2.416 (4.87)
Free flow (SC-Ref)	.0279 (5.73)		.0186 (2.8)	

Slowed down time (SC-Ref)	.0100 (1.92)		.0074 (1.3)	
Running cost (SC-Ref)	.0466 (1.32)		.0516 (1.22))	
Toll cost (SC-Ref)	.0196 (.85)		.0028 (.13)	
Free flow (SC-Ref) -Gains		.0326 (4.2)		.0276 (2.78)
Slowed down time (SC-Ref) - Gains		-.0075 (-.76)		-.0121 (-1.16)
Running cost (SC-Ref) - Gains		.0895 (1.63)		-.0005 (-.01)
Toll cost (SC-Ref) - Gains		.0175 (.45)		-.0307 (-.74)
Free flow (SC-Ref) - Losses		.0198 (2.71)		.0163 (2.02)
Slowed down time (SC-Ref) - Losses		.02169 (2.43)		.0171 (1.92)
Running cost (SC-Ref) - Losses		.1119 (1.47)		.0729 (.86)
Toll cost (SC-Ref) - Losses		.0248 (.80)		.0036 (.10)
Non-random parameters:				
Reference alt constant	.1253 (.83)	.4385 (2.02)		
Reference alt constant APR1			.1924 (1.25)	.7205 (2.82)
Reference alt constant APR2			.1572 (.45)	.4284 (1.02)
Reference alt constant APR3			.3236 (.37)	.5386 (.62)
Reference alt constant APR4			.9293 (1.26)	2.749 (5.0)
SC alt 2 constant	.2006 (2.13)	.2472 (2.34)		
SC constant APR1			.2395 (2.5)	.2243 (2.16)
SC constant APR2			.2839 (1.4)	.3243 (1.53)
SC constant APR3			-.9050 (-1.90)	-.9490 (-2.3)
SC constant APR4			.2935 (.90)	.2825 (1.13)
Toll-specific constant APR2			.2021 (.25)	.4914 (.73)
Toll-specific constant APR3			-1.011 (1.47)	-.8660 (-1.34)
Toll-specific constant APR4			.0132 (.01)	-.5318 (-.68)
Error Component (Alt specific heterogeneity):				
Reference Alternative	2.4874 (17.8)	2.4179 (16.9)		
SC Alternative 1	.5649 (5.58)	.6431 (6.65)		
SC Alternative 2	.1498 (.93)	.3414 (1.57)		
Reference alts (1,2,3,4)			.9050 (7.59)	.5684 (3.56)
SC alt APR1			2.437 (13.9)	2.361 (11.6)
SC alt APR2			1.802 (4.71)	1.828 (4.35)
SC alt APR3			1.123 (2.28)	2.046 (2.84)
SC alt APR4			4.521 (3.40)	3.196 (5.15)
Log-likelihood at convergence	-1716.7	-1714.8	-1715.18	-1708.27

Note: * = constrained triangular random parameter.

Behavioural contrasts are best made through willingness to pay estimates for specific attributes such as values of travel time savings (VTTS). The VTTS summarised in Table 6 are based on conditional distributions (i.e., conditional on the alternative chosen). In the APR models, the VTTS are derived from the subset of rules where there is at least one travel time and one travel cost. There is no VTTS from APR3 and APR4 because there are, respectively, no travel cost and travel time attributes. This does not mean that such people do not have a VTTS, but only that it cannot be revealed from the choice experiment because the context has resulted in the exclusion of relevant attributes.

There is clear evidence that asymmetry exists, with individuals placing a higher value on travel time savings where the non-experienced alternative (i.e. a stated choice alternative pivoted off of the reference alternative) involves a loss of travel time in contrast to a gain in travel time. Gains and losses are weighted differently, acting risk-seeking for losses (i.e., higher WTP) and risk-averse for gains (i.e., lower WTP)[12].

There are a large number of VTTS estimates in Table 6, but the most informative values are the overall weighted averages for free flow and slowed down time, where the weights are the respective running and toll costs. What we see is a substantially higher

12 This result holds for almost all attributes, the one exception being slowed down time under APR2.

mean VTTS for both free flow and slowed down time when comparing models that have common elements in respect of no referencing, symmetrical and asymmetrical referencing (respectively comparisons of columns 2 vs.5; 3 vs. 6; and 4 vs. 7. For free flow time, the models that are based on heterogeneous attribute processing rules give VTTS that are 20-30 percent lower than those where such rules are not accounted for. The VTTS for slowed down time also are much higher when the APR is not allowed for except under non-asymmetric referencing. Overall, our preference is to promote the model that accounts for heterogeneous APR and asymmetric referencing. This model also has the lowest overall log-likelihood value.

Table 6 Valuation of Travel Time Savings (VTTS) Evidence for Car Non-Commuters ($/person hour)

Attribute	Ignoring Process Rule:			Accounting for Process Rule		
	Generic RP time and cost	Non-asymmetric referencing	Asymmetric referencing (gains and losses)	APR-Specific RP Time and Cost	APR-Specific Non-asymmetric referencing	APR-Specific Asymmetric referencing (gains and losses)
Based on running cost:						
Free flow time	15.93 (8.23)	15.02 (10.1)				
Free flow time APR1				11.06 (7.2)	10.8 (4.81)	
Slowed down time	15.89 (6.95)	16.03 (8.37)				
Slowed down time APR1				14.47 (8.51)	14.46 (5.25)	
Free flow time - Gain			14.69 (13.5)			
Free flow time - Loss			16.48 (13.8)			
Slowed down time - Gain			16.27 (9.7)			
Slowed down time - Loss			13.71 (8.1)			
Free flow time – Gain APR1						10.7 (5.34)
Free flow time – Loss APR1						13.02 (6.87)
Slowed down time – Gain APR1						14.26 (4.4)
Slowed down time – Loss APR1						-.73 (0.18)
Based on toll cost:						
Free flow time	9.04 (8.95)	7.56 (5.05)				
Free flow time APR1				9.2 (7.0)	6.83 (4.41)	
Free flow time APR2				19.28 (11.24)	13.69 (10.16)	
Slowed down time	9.00 (6.6	8.0 (5.29)				
Slowed down time APR1				7.88 (5.41)	9.28 (7.41)	
Slowed down time APR2				14.6 (6.33)	7.77 (3.26)	
Free flow time - Gain			8.43 (12.6)			
Free flow time - Loss			9.52 (16.3)			
Slowed down time - Gain			9.19 (8.4)			
Slowed down time - Loss			7.72 (8.0)			
Free flow time – Gain APR1						6.17 (3.8)
Free flow time – Loss APR1						6.89 (4.26)
Free flow time – Gain APR2						11.18 (12.86)
Free flow time – Loss APR2						12.64 (16.1)
Slowed down time – Gain APR1						8.33 (5.09)
Slowed down time – Loss APR1						7.18 (4.73)
Slowed down time – Gain APR2						10.08 (6.39)
Slowed down time – Loss APR2						8.57 (6.90)
Average across Gains and Losses based on running cost						
Free flow time			15.58			
Slowed down time			14.99			
Free flow time APR1						11.86
Slowed down time APR1						6.76
Average across Gains and Losses based on toll cost						
Free flow time			8.98			
Slowed down time			8.45			
Free flow time APR1						6.53
Slowed down time APR1						7.75
Free flow time APR2						11.91
Slowed down time APR2						9.33
Based on weighted average of running and toll cost:						
Free flow time	*12.68*	*11.5*	*12.47*	*10.18*	*8.93*	*10.10*
Slowed down time	*12..63*	*12.2*	*11.91*	*11.36*	*12.02*	*7.95*
Ratio Non-APR: APR VTTS:	Cols 2 vs. 5	Cols 3 vs. 6	Cols 4 vs. 7			
Free flow time	1.245	1.29	1.23			
Slowed down time	1.112	1.02	1.50			

Conclusions

This paper has developed a mixed logit model that incorporates the endogenous decision by individuals to invoke a stage one editing rule in the way that they process pre-defined attributes offered in a stated choice experiment. In addition, recognition of the special role that experienced alternatives play in choice making is incorporated through symmetric and asymmetric referencing.

The model specification has accommodated observed and unobserved heterogeneity through random parameters and error components, the later accommodating additional sources of unobserved heterogeneity for each alternative that are not explained through specific attributes. The approach recognizes the stochastic nature of attribute processing rules which are revealed to the analyst through specific questions in a survey instrument, yet by definition carry a degree of uncertainty in the treatment of specific attributes, which we capture through the endogenous definition of the joint APR and outcome choice.

This study is one contribution to the development of discrete choice models that can simultaneously integrate process and outcome in attribute framing in the presence of experienced alternatives. This approach to seeking out improved ways of capturing the way in which individuals process stated choice experiments and make outcome choices is consistent with arguments being promoted in behavioural and psychological theories of how individuals make choices in real markets. The challenge in stated choice studies is to bring the approach closer to real market responses. The ideas offered herein are consistent with this objective.

References

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

Bhat, C.R., (2003) Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences, *Transportation Research B*, 37(9), 837-855.

Bhat, C.R. and Castela, S. (2002). A unified mixed logit framework for modeling revealed and stated preferences: formulation and application to congestion pricing analysis in the San Francisco Bay area. *Transportation Research B*, 36, 577–669.

Bonini, N., Tentori, K. and Rumiati, R. (2004) Contingent application of the cancellation editing operation: the role of semantic relatedness between risky outcomes, *Journal of Behavioral Decision Making*, 17, 139-152.

Brownstone, D., Bunch, D.S., Train, K. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. Transportation Research, 34B, 315 338.

Bunch, D.S., Louviere, J.J. and Anderson, D.A (1996) A Comparison of Experimental Design Strategies for Choice-Based Conjoint Analysis with Generic-Attribute Multinomial Logit Models, Working Paper, Graduate School of Management, University of California, Davis.

Cantillo, V., Heydecker, B. and Ortuzar, J. de Dios (2006) A discrete choice model incorporating thresholds for perception in attribute values, *Transportation Research* B, 40 (9), 807-825.

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

Greene, W. H. and Hensher, D.A. (in press 2006) Mixed Logit and Heteroscedastic Control for Random Coefficients and Error Components *Transportation Research E*

Hallahan, K. (1999) Seven models of framing: implications for public relations, *Journal of Public Relations Research*, 11(3), 205-242.

Hensher, D.A. (1983) A sequential attribute dominance model of probabilistic choice, *Transportation Research* A, 17A (3), 215-218.

Hensher, D.A. 2001. Measurement of the valuation of travel time savings. *Journal of Transport Economics and Policy* 35(1):71–98.

Hensher, D.A. (2006) Towards a method to establish comparable values of travel time savings from stated choice experiments with differing design dimensions, *Transportation Research,* 40A (10), 829-840.

Hensher, D.A. (in press) How do respondents handle stated choice experiments? - attribute processing strategies under varying information load, *Journal of Applied Econometrics*

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

Hensher, D.A., Rose, J.M. and Greene, W. H. (2005) *Applied Choice Analysis: A Primer,* Cambridge University Press, Cambridge.

Hess, S., Rose, J. M. and Hensher, D.A. (2006) Asymmetrical preference formation in willingness to pay estimates in discrete choice models, Institute of Transport and Logistics Studies, The University of Sydney, July.

Houston, D. A., Sherman, S. J. and Baker, S. M. (1989). The influence of unique features and direction of comparison on preferences. *Journal of Experimental Social Psychology*, 25, 121–141.

Houston, D. A. and Sherman, S. J. (1995). Cancellation and focus: the role of shared and unique features in the choice process. *Journal of Experimental Social Psychology*, 31, 357–378.

Kahneman, D. and Tversky, A. (1979) Intuitive prediction: biases and corrective procedures. In S. Makridakis and S. C. Wheelwright, Eds., *Studies in the Management Sciences: Forecasting*, 12 (Amsterdam: North Holland).

Kanninen, B. J. (2002) Optimal design for multinomial choice experiments, *Journal of Marketing Research*, 39 (2), 214-217.

Kessel, R., Goos, P. and Vandebroek, M. (2006) Comparing algorithms and criteria for designing Bayesian conjoint choice experiments, *Journal of Marketing Research*.

Johnson, F.R., Kanninen, B. and Bingham, M. (2006) Experimental design for stated choice studies, in *Valuing Environmental Amenities using Stated Choice Studies: A Common Sense Approach to Theory and Practice*, Barbara Kanninen, ed., Springer, Dordrecht, The Netherlands, 159-202.

McFadden, D.L. (1981) Econometric models of probabilistic choice in *Structural Analysis of Discrete Data*, Manski, C.F. and McFadden, D.L. (eds.) MIT Press, Cambridge Massachusetts, 198-271.

O'Shaughnessy, J. and O'Shaughnessy, N. (2005) Considerations of equity in marketing and Nozick's decision-value model, *Academy of Marketing Science Review*, 10.

Prelec, D. (1998) The probability weighting function, *Econometrica*, 66, 497-527.

Revelt, D. and K. Train (1998) Mixed Logit with repeated choices: households' choices of appliance efficiency level, *Review of Economics and Statistics*, 80, 1-11.

Rose, J.M., Bliemer, M.C.J., Hensher, D.A. and Collins, A.T. (2005) Designing efficient stated choice experiments involving respondent based reference alternatives, University of Sydney, (under revision for *Transportation Research B*, May 2006).

Sandor, Z. and Wedel, M. (2002) Profile Construction in Experimental Choice Designs for Mixed Logit Models, *Marketing Science*, 21(4), 455–475.

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.

Swait, J. (2001) A non-compensatory choice model incorporating attribute cut-offs, *Transportation Research* B, 35(10), 903-928.

Swait, J. and Adamowicz, W. (2001a) 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. and Adamowicz, W. (2001b) 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.

Thaler, R. (1999) Mental accounting matters. *Journal of Behavioral Decision Making*, 12, 183–206.

Tversky, A. and Kahneman, D. (1981) The framing of decisions and the psychology of choice. *Science*, 211, 453–458.

Tversky, A. and Kahneman, D. (1986) Rational choice and the framing of decisions. *Journal of Business*, 59, 521–278.

Train, K. (2001) A comparison of hierarchical Bayes and maximum simulated likelihood for mixed logit, working paper, Department of Economics, University of California, Berkeley, http://elsa.berkeley.edu/~train/compare.pdf

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