ITLS

WORKING PAPER
ITLS-WP-05-10

The implications of respondent attribute processing rules and experimental design on WTP in stated choice experiments

By

David A Hensher

June 2005

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

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-05-10

TITLE: **The implications of respondent attribute processing rules and experimental design on WTP in stated choice experiments**

ABSTRACT: Individual's process the information in stated choice (SC) experiments in many different ways. In order to accommodate decisions rules that are used in processing information, there is good sense in conditioning the parameterisation of stated choice design attributes on these rules. In particular, rules might be invoked to cope with the dimensionality of the SC design. In this paper we investigate the impact of rules such as attribute aggregation and reference dependency on preference profiles for specific design attributes, as well as the design specification, as we vary the dimensionality of an SC design. Prior to identifying the empirical differences in valuation of travel time savings due to design specification, we account for scale differences from pooling 16 stated choice designs. The heteroscedastic extreme value logit model is estimated to identify the role of design dimensionality and attribute processing rules, after accounting for scale differences across the pooled data designs The empirical evidence, drawn from a study in Sydney of car commuter route choice, suggests that accounting for the way that stated choice designs are processed, given their dimensionality, does make a statistically significant difference on measures of willingness to pay, as does accounting for scale differences between pooled data designs. This evidence has value in guiding the design of SC experiments and in adjusting results from different SC designs when comparing the evidence.

KEY WORDS: ***Stated choice designs, information processing, preference differences***

AUTHORS: 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: June 2005

1. Introduction

Individual's process the information in stated choice experiments in many different ways, in part as a response to the demands of the stated choice design and in part because of the relevancy of information in choice making. To accommodate rules that are used in processing information, there is good sense in conditioning the parameterisation of stated choice design attributes on a set of rules. Examples of such rules are 'adding up' attributes where this is feasible (e.g., travel time components – see Starmer and Sugden 1993) and reference dependency such as contrasts of attribute levels in the SC design relative to recent experiences. In addition, the dimensions of the design (e.g., number of attributes and alternatives, attribute range and levels) may impact on the role that specific attributes play in preference revelation.

This paper is a contribution to a body of research centred on understanding the influence of the survey instrument in the revelation of the preferences of a sample of individuals when faced with evaluating a stated choice experiment and selecting their most preferred alternative. Evidence is accumulating to support trends in key behavioural outputs, such as willingness to pay, that can be attributed to systematic variations in the dimensions of the SC experiment.

Given that we are pooling data obtained for 16 SC designs, we have to account for scale differences if we are to identify the true influence of processing rules and design dimensionality on preferences and hence willingness to pay. The heteroscedastic extreme value (HEV) logit model is selected because it can reveal scale differences through unconstrained variances on the random components associated with each alternative (which are linked to specific SC data designs)[1]. Details of this model are given in Louviere *et al.* (2000).

In this paper, we detail the design of a stated choice experiment, and present the results for the preferred models for commuter choice of a package of route-based trip attributes. We contrast four empirical models in which we exclude/include the scale differences associated with 16 SC designs that are pooled in model estimation, as well as models in which we interact design attributes with two attribute processing rules and design dimensionality. We establish the extent of parameter shift and the implications on the valuation of travel time savings. The empirical evidence suggests that accounting for the way that stated choice designs are processed, given their dimensionality, does make a statistically significant difference on the profile of preferences for specific attributes and alternatives and hence the willingness to pay for travel time savings. Simple adjustments when comparing VTTS from designs with different dimensions such as number of levels for an attribute, the range of levels offered and, in the case of designs pivoting off of a reference alternative are suggested.

[1] It is possible to reveal scale differences using a mixed logit model through having alternative-specific constants specified as random parameters and made a function of a series of dummy variables, where each dummy variable represents a data set. For a pooled data set of 16 designs, used herein, such a model is extremely complex and without a much larger data set, we are unable to estimate a model to control for scale.

2. Individual-specific attribute processing

Individuals bring to the evaluation of a stated choice study a set of attribute processing rules that incorporate the processing and selection rules learnt through choice experience accumulated (and discounted) from the past. Processing rules are typically drawn on to accommodate relevance and complexity (Hensher 2004, in press; Starmer 2000; Swait and Adamowicz 2001a,b; Malhotra, 1982). They include the use of reference dependency (i.e., framing, see Rolfe et al. 2001) as a way of establishing relative net benefit of 'new' alternatives or attributes packages, attribute preservation or elimination (including subtleties of inattention due to irrelevance or cognitive burden), and consequentiality (i.e., questions that have associated with them real reasons for the individual to treat them as of consequence – see Carson et al. 2003). Assumptions that all individuals use the same attribute processing ruls (APR's) when evaluating stated choice experiment treatments run the real risk of imposing substantial biases on parameter estimates in choice models (see Hensher 2004 for some evidence)[2]. The variability in processing is often defined by constructs such as habit formation (e.g., Aarts and Dijksterhuis 2000, Aarts *et al*. 1997) and variety seeking (e.g., Khan 1995), 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 APR.[3]

Arentze *et al.* (2003) scrutinised the influence of task complexity in terms of the number of attributes, alternatives and choice sets presented, as well as the influence of presentation format (surveys with or without pictorial material) including the effects of considering a less literate population. They found that both the presentation method and the literacy level had no significant impacts, while task complexity had a significant effect on data quality.

SC designs have in the main assumed that all attributes are processed in what DeShazo and Fermo (2004) describe as the *passive bounded rationality* model wherein they attend to all information in the choice set but increasingly make mistakes in processing that information. Contrasting this is the *rationally-adaptive* model that assumes individuals recognise that their limited cognition has positive opportunity costs. As DeShazo and Fermo state: "Individuals will therefore allocate their attention across alternative-attribute information within a choice set in a rationally-adaptive manner by seeking to minimise the cost and maximise the benefit of information evaluation" (page 3).

In recognition of the many ways that stated choice experiments are processed, it is important to condition the preferences for attributes and alternatives, revealed within the experimental setting, by the dimensionality of the SC design and other rules acquired by individuals to assist them in any choice setting. To test for a number of processing rules, we have developed a series of stated choice designs, which now detail.

---

[2] Including false assumptions about lexicographic choice behaviour.

[3] Such a 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.

3. The design plan

The data are drawn from a larger study reported in Hensher (in press, 2004) in which 16 stated choice designs (Table 1) have been developed. A sample of car commuters travelling in Sydney in 2002 defines the application context. The data was collected specifically to investigate the influence of different SC designs on preference revelation; however useful policy outputs on values of travel time savings can be obtained and used to evaluate the benefits of tollroads, which in large measure deliver substantial time savings over alternative non-tolled routes.

Each commuter evaluated one of the 16 designs. Across the full set of stated choice experiments, the designs differed in terms of the number, range and levels of attributes, the number of alternatives and the number of choice sets. The combination of the dimensions of each design is often seen as the source of design complexity (Dellaert et.,al. 1999) and it is within this setting that we have varied the number of attributes that each respondent is asked to evaluate. The overall sample was built up by having an inbuilt random number generator that selected one of the designs each time a respondent is interviewed.

Table 1 Dimensions of each Design

Number of choice sets	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

The candidate attributes have been selected based on earlier studies reported mainly in the marketing and transportation literature (see Hensher in press, Ohler et al. 2000). They are: free flow time (FFT), slowed down time (SDT), stop/start time (SST), trip time variability (TTV), toll cost (TLC), and 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:

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

The specific SC design is three unlabelled alternatives that have attribute levels that pivot off the levels associated with a current car-commuting trip. That is, the *actual* levels of the attributes shown to respondents are calculated relative to those of the experienced reference car commuter trip. The levels applied to the choice task (see Appendix B) differ depending on the range of attribute levels and the number of levels for each attribute.

The designs are computer-generated. They aim at minimising the correlations between attributes and maximising the amount of information captured by each choice set. We maximised the determinant of the covariance matrix, which is itself a function of the estimated attribute parameters (within the experimental design literature this is known as D-optimality – see appendix for further details). The design dimensions are translated into SC screens as illustrated in Figure 1.

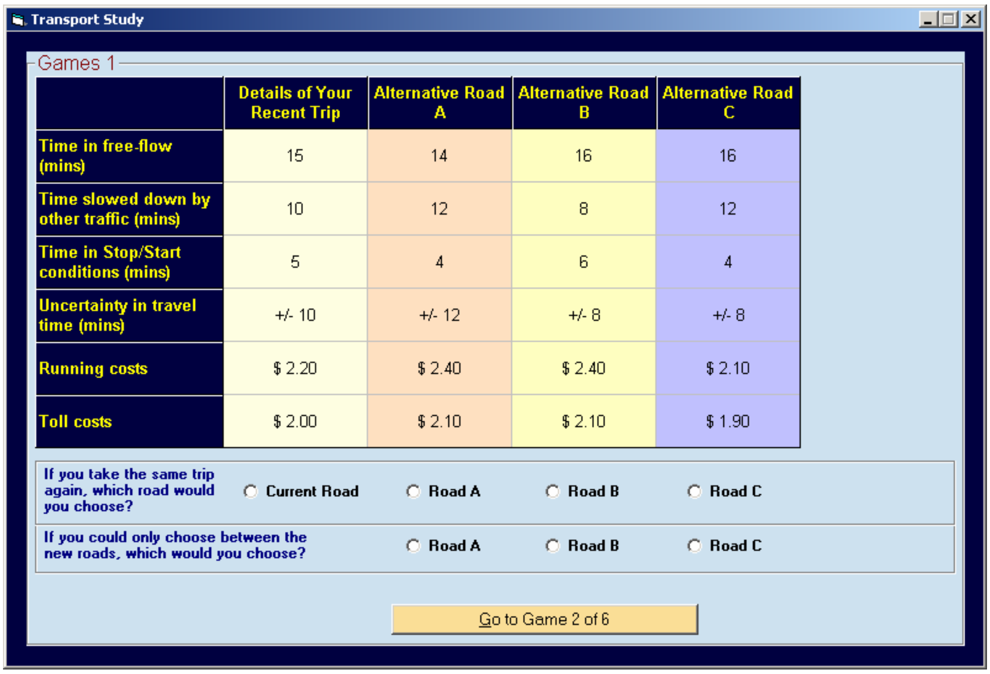

Figure 1 Example of a stated choice screen

4. Results

Four model specifications have been estimated:

M1 – Multinomial logit without the APR’s and design dimensions but with naïve pooling (i.e. ignore scale differences) of SC data sets
M2 - Multinomial logit with APR’s, design dimensions and naïve pooling
M3 – HEV logit without the APR’s and design dimensions but accounting for scale differences of SC data sets
M4 – HEV logit with APR’s and design dimensions and accounting for scale differences of SC data sets

Two attribute processing rules were investigated – the extent to which individuals first add up any attributes such as travel time components before assessing the alternatives; and the absolute differences between an attribute’s level reported for a recent car commuter trip and the level presented in an SC design generated alternative. These conditions are interacted with each of the travel time attributes.

The additivity of attributes in the processing of the choice sets was identified from a series of supplementary questions. The evidence, summarised in Table 2, suggests that there is a substantial amount of aggregation of the travel time and cost components in evaluation of the alternatives, with over 75% of the respondent’s aggregating all travel time dimensions (who did not face a single total time attribute) as part of the way they process the attribute information. Aggregation of cost was also high, although it was not found to be statistically significant interactive influence on the marginal (dis)utility of cost.

Table 2 Summary of Attribute Role and Treatment of Additivity in Respondent’s Processing of SC Screens (proportion of relevant observations). Blank cells mean not applicable.

	Proportion of sample who added up components of:	
Design	Time	Cost
0	.781	
1	.794	
2	.829	
3	.853	
4		
5		
6	.758	.636
7		
8	.839	.613
9	.871	.677
10	.793	
11	.900	
12	.800	.760
13		
14	.893	
15	.750	

The final models (M1-M4) are summarised in Table 3. The overall statistical fit of the all models is impressive (with pseudo-R^2's greater that 0.65). The likelihood ratio test for M2 vs. M1, and for M4 vs. M3, with a difference of eight degrees of freedom, rejects at the 95% confidence interval, the null hypothesis of no statistical difference.

Table 3 Final Model Results
Time is in minutes, cost is in dollars; t-values in brackets, 4,593 observations.

Attribute	M1 (no scaling)	M2 (no scaling)	M3 (scaling)	M4 (scaling)
	Parameter estimates (t-values)	Parameter estimates (t-values)	Parameter estimates (t-values)	Parameter estimates (t-values)
Free flow time	-0.1271 (-19.64)	-.0176 (-5.99)	-0.1856 (-7.20)	-0.1786 (-3.67)
Non-free flow time*	-0.1203 (-23.52)	-0.1549 (-6.66)	-0.1757 (-7.44)	-0.1638 (-3.86)
Total time	-0.1671 (-20.52)	-0.2957 (-13.03)	-0.1772 (-18.1)	-0.2933 (-12.92)
Running cost	-0.7217 (-8.18)	-0.7270 (-8.22)	-1.1118 (-5.37)	-0.7685 (-4.10)
Toll cost	-1.1009 (-8.95)	-1.1061 (-8.95)	-1.7459 (-5.63)	-1.1854 (-4.24)
Total cost	-0.8809 (-15.14)	-0.7619 (-13.30)	-1.1420 (-9.03)	-0.7583 (-5.48)
Total time	-0.1671 (-20.52)	-0.2957 (-13.03)	-0.1772 (-18.12)	-0.2933 (-12.92)
APR interactions:				
Free flow time x add time dummy		-0.0382 (-2.42)		-0.0380 (-2.20)
Free flow time x reference dependency		-0.0004 (-1.96)		-0.0006 (-2.10)
Non-free flow time x add time dummy		-0.0347 (-3.12)		-0.0349 (-2.70)
Design Dimension interactions:				
Free flow time x # levels of attribute		0.0176 (2.52)		-0.0177 (2.10)
Free flow time x wide attribute range in design		0.0348 (2.82)		0.0255 (1.90)
Non-free flow time x # levels of attribute		0.0162 (2.56)		0.0181 (2.31)
Non-free flow time x wide attribute range in design		0.0312 (3.26)		0.0294 (2.46)
Total time x wide attribute range in design		0.1649 (6.84)		0.1634 (6.59)
Data Design Scale parameters:			1.0,0.77,0.65,0.61,0.98,0.77,0.65,0.60,1.04, 0.79,0.66,0.62,1.01,0.76,0.66,0.62#	1.0,1.15,0.92,0.88,0.99,1.16,0.93,0.87,1.04, 1.19,0.95,0.91,1.01,1.15,0.95,0.89#
Log-Likelihood	-3669.66	-3616.43	-3646.59	-3594.72

* Non-free flow time is the sum of all other time components (i.e. slowed down and stop-start time)
# All scale parameters are highly significant with all t-ratios greater than 6.4 for M3 and greater than 4.3 for M4.

The values of travel time savings for each model are summarised in Table 4. Models 2 and 4 permit the derivation of distributions of VTTS across the sampled population, due to the influence of specific attribute processing rules and the dimensionality of the design set. These are graphed in Figures 2 and 3. The differences between the scaled and non-scaled models are visually small for each travel time attribute; however a statistical test of differences between the VTTS distributions finds, for all three attributes, that we can reject the null hypothesis of no differences in all cases (at the 95% percent confidence level). The z-values are respectively 5.35, 5.9 and 7.79 for free flow, non-free flow and total time. Hence failure to account for differences in scale does impact on the VTTS, and for this specific application, it biases the estimates upwards, on average by 1.9%, 2.38% and 6.55% for each of free flow, non-free flow and total time.

The evidence suggests that attribute processing, by adding up components of travel time and contrasting the SC attribute level with the reference alternative's attribute level, have a statistically significant influence on preferences. Specifically, after controlling for scale differences across the data sets (M4), individuals who add up the time components tend to have higher marginal disutility of travel time than individuals who keep the components separate[4]. Another way of stating this is that individuals who aggregate the time components as an attribute processing rule, have a higher VTTS, all other things being equal. The same directional impact occurs for individuals who compare an SC alternative's attribute level against a reference alternative. When an individual is faced with a greater difference between the reference and SC attribute level, the VTTS will increase, all other things being equal;

When we account for possible interactions between attributes and design specifications, we find for free flow time, that where an individual evaluates alternatives based on a higher number of levels used and a wider range (relative to narrow and a base) of attribute levels in the design, the VTTS tends to decrease, all other things being held constant. For non-free flow time, the same directional implications apply. The evidence on attribute range confirms the findings of a number of other studies (e.g., Ohler *et al.* 2000, Hensher 2004); namely a design with a relatively narrow attribute range will increase VTTS relative to a wider range. This type of evidence is potentially very worrying, since one interpretation is that the analyst can create an appropriate VTTS solution simply by the way they design their SC experiment.

Table 4 VTTS Findings* ($ per person hour) standard deviation across sample in brackets for models M2 and M4

	M1	M2	M3	M4
Free flow time	10.57	12.07 (2.24)	10.01	11.84 (1.97)
Non-free flow time	10.01	10.33 (1.98)	9.48	10.09 (1.92)
Total time	13.89	17.23 (6.79)		16.17 (6.37)

* Based on parameter of running cost.

[4] The application of such a rule does not necessarily suggest that the components are not distinguished, but rather that the components are assessed in the context of the total trip time. This distinction is subtle and potentially complex, requiring further research to understand exactly what is being processed. With 81 percent of the sample indicating that it evaluated the component of time within the context of adding them up, then this is an important issue to resolve. It should not however be assumed that future SC designs should simply offer a total travel time attribute, until we have convincing evidence that the differences in the marginal (dis)utilities associated with the components do not matter.

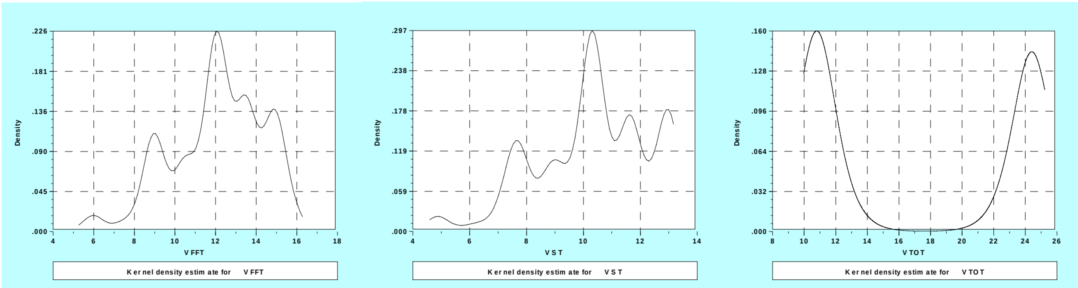

Figure 2 Distributions of VTTS for Model 2

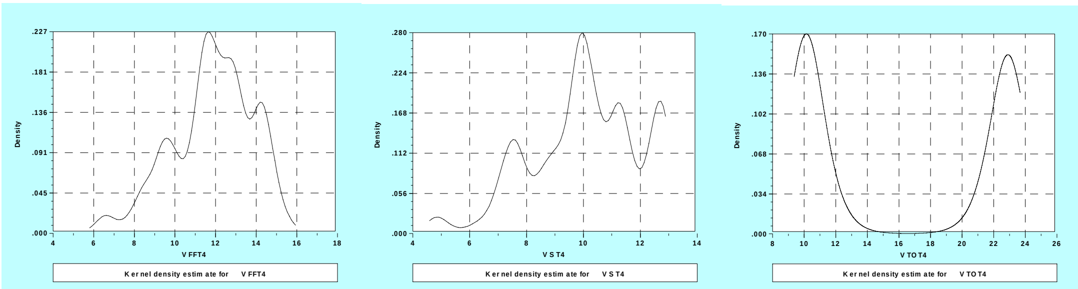

Figure 3 Distributions of VTTS for Model 4

5. Conclusions

This paper is a contribution to a body of research centred on understanding the influence of the survey instrument in the revelation of the preferences of a sample of individuals when faced with evaluating a stated choice experiment and selecting their most preferred alternative. Evidence is accumulating to support trends in key behavioural outputs, such as willingness to pay, that can be attributed to systematic variations in the dimensions of the SC experiment.

These design dimensions induce (in part at least) specific processing rules as mechanisms for coping with the specification of the design (both quantitatively and qualitatively). However, and importantly, the behavioural responses may be associated with processing rules that individuals use in many circumstances that are not unique to processing SC experiments, and which are brought to bear on the SC task in hand.

Two important empirical findings on attribute processing are that (i) SC designs in which the attribute levels deviate less from the reference (or experienced) level, are more likely to produce lower mean VTTS than those where the difference is greater; and (ii) where an attribute has components that are potentially additive (as in components of travel time), the mean VTTS is higher when a respondent evaluates the components via an addition rule.

The important behavioural inferences that can be drawn from consideration of the dimensionality of a design are that lower (relative) mean estimates of VTTS appear to be associated with designs that have a wider range on each attribute, and a greater number of levels per attribute. The differences cannot be used to conclude that specific designs are ‘better’ than other designs in a relevancy sense. But they do send a very strong warning about comparing outputs from different stated choice studies.

Appendix A
D-Efficient (Optimal) Designs

Traditional designs, such as orthogonal designs, ensure that we can estimate the effects of the different attributes independently of each other. In contrast, a D-optimal design considers explicitly the importance of the levels of the attributes, and ensures that the alternatives in the choice sets provide more information about the trade-offs between the different attributes. However, this requires explicit incorporation, in the design, of prior information about the respondents' preferences.[5] Possible sources of information include results obtained from previous studies, or information obtained from pilot studies.

Optimal designs will be statistically efficient but will likely have correlations, orthogonal fractional factorial designs will have no correlations but may not be the most statistically efficient design available. Hence, the type of design generated reflects the belief of analysts as to what is the most important property of the constructed design. Carlsson and Martinsson (2003) have recently shown, using Monte-Carlo simulation, that D-optimal designs, like orthogonal designs, produce unbiased parameter estimates but that the former have lower mean (see also Bliemer and Rose 2005).

In determining the most statistically efficient design, the literature has tended towards designs which maximise the determinant of the variance-covariance matrix, otherwise known as the Fisher information matrix, of the model to be estimated. Such designs are known as D-optimal designs. In determining the D-optimal design, it is usual to use the inversely related measure to calculate the level of D-efficiency, that is, minimise the determinant of the inverse of the variance-covariance matrix. The determinant of the inverse of the variance-covariance matrix is known as D-error and will yield the same results maximising the determinant of the variance-covariance matrix.

Using the multinomial logit (MNL) model as an example (but recognising that the design will vary according the choice of discrete choice model), the log likelihood function of the MNL model is shown as equation (A1).

$$L = \sum_{n=1}^{N} \sum_{s=1}^{S} \sum_{j=1}^{J} y_{njs} \ln(P_{njs}) + c \tag{A1}$$

where y_{njs} is a column matrix where 1 indicates that an alternative j was chosen by respondent n in choice situation s and 0 otherwise, P_{njs} represents the choice probability from the choice model, and c is a constant. Maximising equation (A1) yields the maximum likelihood estimator, $\hat{\beta}$, of the specified choice model given a particular set of choice data. McFadden (1974) showed that the distribution of $\hat{\beta}$ is asymptotically normal with a mean, β, and covariance matrix

[5] Orthogonal designs also require prior information in order to choose the attribute levels in such a way that dominating and inferior attributes are avoided.

$$\Omega = \left(X'PX\right) = \left[\sum_{m=1}^{M}\sum_{j=1}^{J} x'_{njs}\; P_{njs} x_{njs}\right] \quad \text{(A2)}$$

and inverse,

$$\Omega^{-1} = \left(X'PX\right)^{-1} = \left[\sum_{m=1}^{M}\sum_{j=1}^{J} x'_{njs}\; P_{njs} x_{njs}\right]^{-1}. \quad \text{(A3)}$$

where P is a *JS*×*JS* diagonal matrix with elements equal to the choice probabilities of the alternatives, *j* over choice sets, *s*. For Ω, several established summary measures of error have been shown to be useful when contrasting designs. The most popular summary measure is known as D-error, inversely related to D-efficiency.

$$\text{D-error} = \left(\det \Omega^{-1}\right)^{\frac{1}{K}} \quad \text{(A4)}$$

where *K* is the total number of *generic* parameters to be estimated from the design. Minimisation of equation (A4) will produce the design with the smallest possible errors around the estimated parameters. Kanninen (2002) and Kuhfeld *et al.* (1994) provide further details.

Appendix B

Attribute Profiles for the Entire Design

(units = %)	**Base range**			**Wider range**			**Narrower range**		
Levels:	2	3	4	2	3	4	2	3	4
Free flow time	± 20	-20, 0, +20	-20,-10,+10,+20	-20, +40	-20,+10,+40	-20, 0,+20,+40	± 5	-5, 0,+5	-5, -2.5, +2.5, +5
Slow down time	± 40	-40, 0, +40	-40,-20,+20,+40	-30, +60	-30,+15,+60	-30, 0,+30,+60	± 20	-20, 0, +20	-20, -2.5, +2.5, +20
Stop/start time	± 40	-40, 0, +40	-40,-20,+20,+40	-30, +60	-30,+15,+60	-30, 0,+30,+60	± 20	-20, 0, +20	-20, -2.5, +2.5, +20
Slow down-stop/start time	± 40	-40, 0, +40	-40,-20,+20,+40	-30, +60	-30,+15,+60	-30, 0,+30,+60	± 20	-20, 0, +20	-20, -2.5, +2.5, +20
Total travel time	± 40	-40, 0, +40	-40,-20,+20,+40	-30, +60	-30,+15,+60	-30, 0,+30,+60	± 20	-20, 0, +20	-20, -2.5, +2.5, +20
Uncertainty of travel time	± 40	-40, 0, +40	-40,-20,+20,+40	-30, +60	-30,+15,+60	-30, 0,+30,+60	± 20	-20, 0, +20	-20, -2.5, +2.5, +20
Running costs	± 20	-20, 0, +20	-20,-10,+10,+20	-20, +40	-20,+10,+40	-20, 0,+20,+40	± 5	-5, 0,+5	-5, -2.5, +2.5, +5
Toll costs	± 20	-20, 0, +20	-20,-10,+10,+20	-20, +40	-20,+10,+40	-20, 0,+20,+40	± 5	-5, 0,+5	-5, -2.5, +2.5, +5
Total costs	± 20	-20, 0, +20	-20,-10,+10,+20	-20, +40	-20,+10,+40	-20, 0,+20,+40	± 5	-5, 0,+5	-5, -2.5, +2.5, +5

References

Aarts, H. and Dijksterhuis, A. (2000) The automatic activation of goal-directed behaviour: the case of travel habit, *Journal of Environmental Psychology*, 20, 75-82.

Aarts, H., Verplanken, B. and van Knippenberg, A. (1997) Habit and information use in travel mode choices, *Acta Psychologica*, 96, 1-14.

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

Bliemer, M.C.J and Rose, J.M. (2005) Efficient Designs for Alternative Specific Choice Experiments, Institute of Transport and Logistics Working Paper, University of Sydney, February.

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

Carson, R.T., Groves, T. and Machina, M. (2003) Incentive and informational properties of preference questions, Department of Economics, University of California San Diego, March.

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 (2004) Implications of rationally-adaptive pre-choice behaviour for the design and estimation of choice models, working paper, School of Public Policy and Social Research, University of California at Los Angeles, August.

Gilboa, I. and Schmeidler, D. (2001) *A Theory of Case-Based Decisions*, Cambridge University Press, Cambridge.

Hensher, D.A. (2004) Accounting for Stated Choice Design Dimensionality in Willingness to Pay for Travel Time Savings, *Journal of Transport Economics and Policy*, 38 (2), 425-446.

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

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

Kanninen, B.J. (2002) Optimal Design for Multinomial Choice Experiments, *Journal of Marketing Research*, 39 (May), 214-217.

Khan, B.E. (1995) Consumer variety-seeking among goods and services: An integrative review, *Journal of Retailing and Consumer Services*, 2 (3), 139-148.

Kuhfeld, W.F., Tobias, R.D., and Garratt, M. (1994) Efficient Experimental Design with Marketing Research Applications, Journal of Marketing Research, 21 (November), 545-557.

Louviere, J.J., Hensher, D.A. and Swait, J.F. (2000) *Stated Choice Methods and Analysis*, Cambridge University Press, Cambridge.

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

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

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

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

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

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.