WORKING PAPER

ITLS-WP-10-21

Interrogation of responses to stated choice experiments: Is there sense in what respondents tell us?

By

David A Hensher and Andrew T Collins

November 2010

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-10-21

TITLE: **Interrogation of responses to stated choice experiments: Is there sense in what respondents tell us?**

ABSTRACT: There is an extensive and growing literature on the design and use of stated choice experiments. Such experiments are generally regarded as the preferred framework within which to collect data that is able to reveal the preferences of individuals in a particular application context. Analysts have developed sophisticated ways of analysing such data, typically using a form of discrete choice model that identifies the marginal (dis)utility associated with each observed attribute linked to an alternative, as well as accounting for sources of preference and scale heterogeneity. There is also a growing literature studying the attribute processing rules (or heuristics) that respondents use as a way of simplifying the task of choosing, for all manner of reason. We find that there is relatively less effort placed on looking at the data defining each choice situation for each individual, as represented by the stated choice experiment and the choice response, and seeing if there are 'evidential' rules that support in a rational way, the choice responses. This is a complex issue since individuals are complex in their decision making, but it is possible to build up a number of 'evidential' rules that can be used to establish confidence in the stated choices. This paper explores the evidence in building confidence in the data used to infer the many behavioural outputs offered from the estimation of choice models using stated choice data. We find a high level of confidence in the evidence, and identify at least two features of choice set processing, namely value learning and majority of confirming dimensions, that are worthy of inclusion in the estimation of choice models. The evidence suggests that there is a great deal of behavioural sense in stated choice responses, for all manner of possible reason.

KEY WORDS: ***Choice experiments, rational choice, sequence effects, referencing, reference revision, attribute processing, majority of confirming dimensions, evidential rules***

AUTHORS: **David A Hensher and Andrew T Collins**

CONTACT:

Institute of Transport and Logistics Studies (C37)
The Australian Key Centre in Transport Management
The University of Sydney NSW 2006 Australia

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

DATE: November 2010

Acknowledgements. We thank John Rose, Joffre Swait, Sean Puckett and Victor Adamowitz for very useful comments on earlier versions.

1. Introduction

Stated choice experiments are used extensively to create data capable of modelling choices to obtain parameter estimates that describe the preferences of individuals for specific attributes of alternatives within a pre-defined choice setting (Louviere *et al.* 2000). The popularity of such choice experiments is in part a product of the lack of appropriate revealed preference data in situations where choosing amongst a number of alternatives can be observed in real markets, but also due to the ability, within a single unified theoretical framework, to investigate the potential take up of alternatives which do not currently exist in terms of the levels and mix of attributes and/or uniqueness beyond a set of prescribed attributes.

It is common practice for analysts to pool the data from a sample of respondents, accounting for the presence of multiple observations within each respondent, and then to estimate a discrete choice model, accounting to varying degrees for observed and unobserved preference heterogeneity, and more recently also scale heterogeneity (see Fiebig et al. 2009 and Greene and Hensher 2010 as examples). There is also a growing interest in investigating the role that specific attribute processing heuristics play in conditioning the influence of each attribute associated with each alternative (see Hensher 2010 for an overview, Hess and Hensher (in press 2010), and Cameron and DeShazo 2008), using a variety of self-explication and functional-form inferred responses. Another area of growing interest, particularly in the non-market valuation literature, is research into behavioural explanations for the preference changes that appear to occur over a sequence of choice tasks, using parametric (Bateman *et al.* 2008, Day *et al.* 2009, McNair *et al.* 2010a) and non-parametric tests (Day and Pinto 2010) and equality-constrained latent class models (McNair *et al.* 2010b).

What we believe is not given enough emphasis is the extent to which we can learn from an interrogation of each response at the individual and choice set level, and set up candidate rules that align with one or more possible processing rules used by an individual, to reveal their choice response that satisfy specific analyst-prescribed rationality tests. While we can never be certain that the rule is applied, we are seeking out a way to gain confidence in the evidence, given that some pundits believe that respondents are known to make choices which have no rational attachment.

To illustrate the focus of this paper, we reproduce, in Table 1, data from one respondent in one of many choice experiments the authors have conducted, in the context of choosing amongst three routes for a commuter trip, where the first route description is the reference or status quo (SQ) alternative associated with a recent trip. The design attributes are free flow time (FF), slowed down time (SDT), running cost (Cost), toll if applicable (Toll), and overall trip time variability (Var) (times are in minutes, costs in dollars, and time variability in plus or minus minutes). Focussing on these five attributes only, we highlight in green the most attractive level (e.g., lowest FF), and propose that if an alternative had the most attractive level on at least one attribute, and that alternative was chosen, then we can reasonably suggest that the respondent was rational in their choice, based of course on only the offered attributes. There could be other reasons why an alternative is chosen, regardless of the attribute levels and relativity, such as satisfaction with the status quo or the adoption of a minimum regret calculus in contrast to a utility maximisation calculus (see Chorus 2009 and Hensher *et al.* 2010). However, on the face of the observed attribute evidence, the 16 choice scenarios satisfy a rule of rational choice in 16 situations. Five of the choice scenarios show the status quo as the preferred alternative. The ‘rationality’ test assumes that all attributes (and levels) are relevant and that a fully compensatory processing strategy is active. It may be that this individual adopts one or more attribute processing rules in evaluating the choice scenarios, which may be the basis of choice in any of the 16 choice sets, regardless of whether they have passed the ‘rationality’ test used above.

Supplementary data associated with self-explication on whether specific attributes were ignored or added up (where they have a common metric) might also be brought to bear, to add additional insights into the choice responses. No attributes were ignored by this respondent, as reported by the self-explication questions. Looking at the possibility that this individual may have added up FF and SDT and/or COST and TOLL, we cannot find any evidence within the 'rationality' test that it would have failed if attribute addition (TotTime, TotCost) had not been applied, although this may have assisted in making the choice.

Table 1: Example of 16 choice scenarios evaluated by a respondent

Choice scenario	Alternative	TotTime	TotCost	Var	FF	SDT	Cost	Toll	Choice1	Rational = Y
1	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
1	2	48	5.7	8	14	34	2.6	3.1	1	Y
1	3	36	8	6	14	22	4.5	3.5	0	Y
2	1 (SQ)	40	5.4	25	12	28	3.2	2.2	1	Y
2	2	40	7.1	8	6	34	4.5	2.6	0	Y
2	3	44	4.7	6	10	34	1.6	3.1	0	Y
3	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
3	2	28	7	8	14	14	3.5	3.5	1	Y
3	3	40	2.6	6	6	34	2.6	0	0	Y
4	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
4	2	28	4.5	2	14	14	4.5	0	1	Y
4	3	48	4.2	8	14	34	1.6	2.6	0	Y
5	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
5	2	44	8	4	10	34	4.5	3.5	0	Y
5	3	36	1.6	2	14	22	1.6	0	1	Y
6	1 (SQ)	40	5.4	25	12	28	3.2	2.2	1	Y
6	2	48	5.1	6	14	34	1.6	3.5	0	Y
6	3	48	3.5	4	14	34	3.5	0	0	Y
7	1 (SQ)	40	5.4	25	12	28	3.2	2.2	1	Y
7	2	44	6.6	2	10	34	3.5	3.1	0	Y
7	3	48	6.1	8	14	34	2.6	3.5	0	Y
8	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
8	2	36	7.6	6	14	22	4.5	3.1	0	Y
8	3	20	5.1	4	6	14	1.6	3.5	1	Y
9	1 (SQ)	40	5.4	25	12	28	3.2	2.2	1	Y
9	2	48	4.2	2	14	34	1.6	2.6	0	Y
9	3	28	6.6	8	6	22	3.5	3.1	0	Y
10	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
10	2	20	4.7	4	6	14	1.6	3.1	1	Y
10	3	44	7	2	10	34	3.5	3.5	0	Y
11	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
11	2	32	1.6	8	10	22	1.6	0	1	Y
11	3	28	6.1	6	14	14	3.5	2.6	0	Y
12	1 (SQ)	40	5.4	25	12	28	3.2	2.2	1	Y
12	2	48	2.6	4	14	34	2.6	0	0	Y
12	3	40	7.1	2	6	34	4.5	2.6	0	Y
13	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
13	2	24	5.2	6	10	14	2.6	2.6	1	Y
13	3	48	7.6	4	14	34	4.5	3.1	0	Y
14	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
14	2	40	3.5	6	6	34	3.5	0	1	Y
14	3	32	5.2	4	10	22	2.6	2.6	0	Y
15	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
15	2	36	6.1	4	14	22	3.5	2.6	0	Y
15	3	28	5.7	2	14	14	2.6	3.1	1	Y
16	1 (SQ)	40	5.4	25	12	28	3.2	2.2	0	Y
16	2	28	6.1	2	6	22	2.6	3.5	1	Y
16	3	24	4.5	8	10	14	4.5	0	0	Y

The following sections of the paper undertake a more formal inquiry using another data set collected in 2007 in New Zealand. The paper is organised as follows. We briefly describe the data, followed by a statistical assessment of the data in the search for possible rules that explain specific choice responses that are rational or not under specific assumptions. We then discuss the evidence, and conclude with a statement of the degree of confidence one might have in the behavioural sense of the data emanating from a stated choice experiment.

2. The data setting

As part of a larger study to evaluate the costs and benefits of a new tollroad proposal in New Zealand, we undertook field work in late 2007 to identify the preferences of a sample of 132 commuters, 115 non-commuters and 124 individuals travelling on employer business in the catchment area south of Auckland. A stated choice experiment was included together with questions that sought information of a recent trip which was used to construct the pivoted attribute levels of the non-status quo or reference alternatives. There were 16 choice scenarios in which the respondent compared the levels of times and costs of a current/recent trip against two alternative opportunities to complete the same trip described by other levels of times and costs. The respondent had to choose one of these alternatives. The profile of the attribute range is given in Table 2 with an illustrative stated choice scenario screen in Figure 1.

Table 2: Profile of the attribute range in the choice experiment design

Attribute	Levels
Free Flow Time	-0.3, -0.15, 0, 0.15, 0.3
Slowed Down Time	-0.3, -0.15, 0, 0.15, 0.3
Trip Time Variability	±0%, ±,5%, ±10%, ±15%
Running Cost	-0.4, -0.1, 0 ,0.2, 0.4
Toll Cost	\$0, \$0.5, \$1, \$1.5, \$2, \$2.5, \$3, \$3.5, \$4

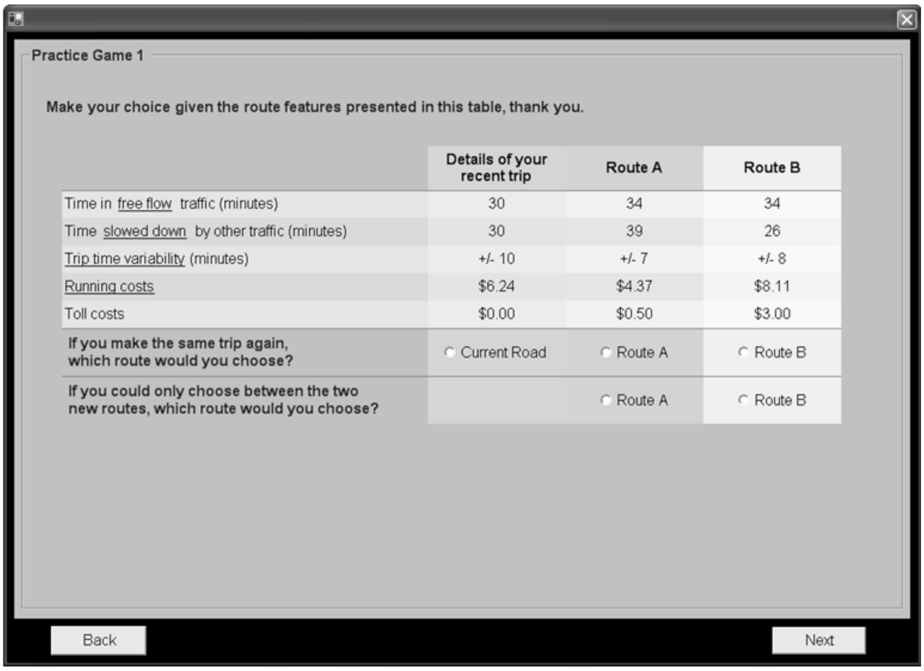

Figure 1: An example of a stated choice screen

A few additional rules were imposed on the design:

(i) Free flow and slowed times[1] were set to five minutes if the respondent entered zero for their current trip;

(ii) Some travel time variability was enforced to ensure that respondents entered a best and worst departure time that was different to their actual leave time; given that it is feasible for their recent trip to have the best or worst departure time. We allowed the best and worst departure times to be reported as the same as their recent time, and where this occurred we set an artificial base as per the same rule in (i).

In addition, supplementary questions were asked upon completion of all 16 choice scenarios on whether specific attributes were ignored. Given the focus of this paper, other details of the study are not provided.

3. Investigating candidate evidential rules

3.1 Overall descriptive profile

The rationality test presented above for one observation can be applied across the 6,048 observations in the New Zealand data. When all attributes are assumed to be relevant, we find that over 99.9 percent of the observations pass the rationality test associated with at least one attribute being best on the chosen alternative. Over 99.7 percent of the observations pass this test when we allow for the self-explicated response on attribute non-preservation.

We ran simple logit models to explore the possible influence of the commuter's age, income and gender under full attribute relevance and attribute self-explication. Income and gender had no influence, but age had a statistically significant impact under attribute processing (based on respondent self-explication of whether an attribute was ignored or not), with the probability of satisfying the rationality test increasing as the commuter ages.

Table 3 shows the proportion (and counts) of rational choice sets by choice task sequence number, suggesting that out of 6,048 choice sets for 378 respondents, there is no noticeable deterioration in rational response as the respondent works through the choice sets from set 1 to set 16[2].

[1] The distinction between free flow and slowed down time is solely 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.

[2] We also ran two simple logit models in which the dependent variable was a binary variable, (where 1 indicated that the alternative chosen had at least one attribute that was best across all three alternatives for a choice scenario), and a series of 15 choice sequence dummy variables, to see if there was a relationship between choice scenario sequence and response rationality. One model assumed all attributes are relevant, and the other accounted for the respondent's self-explicated response on whether an attribute was ignored or not. We were unable to identify any systematic influences under full relevance; in the case of attribute preservation, there were also no significant effects.

Table 3: Influence of choice sequence on choice response

Choice Set Sequence	Assuming full attribute relevance		Allowing for attribute non-preservation	
	Proportion rational	Count irrartional	Proportion rational	Count irrartional
1	0.999339	4	0.996693	20
2	1	0	0.996858	19
3	0.999182	4	0.996197	23
4	0.999339	4	0.99752	15
5	0.999835	1	0.997024	18
6	0.999008	6	0.99752	15
7	0.999504	3	0.999008	6
8	0.999669	2	0.997189	17
9	0.999008	6	0.996528	21
10	0.999339	4	0.997851	13
11	0.999339	4	0.99752	15
12	0.999504	3	0.997189	17
13	0.999339	4	0.997189	17
14	0.999339	4	0.997024	18
15	0.999339	4	0.998181	11
16	1	0	0.998347	10

Choice task response latencies have been used by Haaijer *et al.* (2000) and Rose and Black (2006) to improve the model fit of the final choice models of interest. We took an alternative approach, investigating the relationship between the rationality test (both under full attribute relevance and stated attribute attendance) and both the amount of time to complete the survey, and the amount of time to complete each of the 16 choice scenarios against the rationality test (i.e., the response latency). The only statistically significant relationship, as reported in Table 4, was between both total time and choice scenario completion time and consistency with the rationality test under full attribute relevance, and when attribute processing was taken into account in determining compliance with the rationality test. We find that respondents who satisfied the rationality test, after accounting for attribute processing, tended to spend more time, on average 129.6 seconds overall (noting the sample mean of 1,787 seconds) or 87.5 seconds more on the choice scenarios (noting the sample mean of 703.9 seconds) compared to respondents who failed the rationality test on one or more choice scenarios. The average time was 27.48 seconds, with a standard deviation of 26.03 seconds.

Table 4: Choice scenario completion time influences

Simple Regression

	Full Relevance	**APS Ignore**
Constant	27.478 (142.8)	22.1163 (30.8)
Full Relevance Rationality Test (1,0)	0.0081 (4.39)	
Rationality Test under Attribute Non-Preservation (1,0)	-	5.5856 (7.5)
R-squared	0.00017	0.0019
Sample size	6048	

3.2 Derivative willingness to pay

The next test was to estimate choice models that distinguished between (i) the full sample (6,048 observations) assuming all attributes are relevant (Full), (ii) the full sample with choice scenarios removed when the rationality test failed (5,995 observations) (Rational), (iii) the full sample taking into account a self-explicated attribute processing strategy (6,048 observations) (Full APS), and (iv) the full APS sample with choice scenarios removed when the rationality test failed (5,793 observations) (Rational APS). The findings on values of travel time savings (VTTS) are summarised in Table 5, based on both the running cost (RC) and toll cost (TC) parameter estimates[3]. We have also included the percentage changes in the mean VTTS estimates as a way of identifying the behavioural implications of failing the strict rationality test, as defined by the observed attributes that at least one attribute is the best for the chosen alternative, regardless of whether it was the reference alternative or not.

While the differences are marked in some cases, none of the differences in mean VTTS are statistically different, using the delta test to obtain standard errors[4]. This is the case even when over four percent of the sample is removed due to a suspicion of irrational choice behaviour. This finding suggests that the underlying model is robust, and able to cope with a small percentage of seemingly irrational decisions.

Table 5: Implications of the rationality test on mean value of travel time savings

	Running cost					
	All attributes relevant			Attribute processing strategy applied		
$/person hour (VTTS)	Full	Rational	Difference	Full APS	Rational APS	Difference
Free flow time	$13.01	$12.53	3.81%	$12.02	$11.62	3.51%
Slowed down time	$13.93	$13.85	0.62%	$14.52	$14.53	-0.09%
Trip time variability	$2.57	$2.53	1.51%	$2.33	$2.95	-20.83%
	Toll cost					
	All attributes relevant			Attribute processing strategy applied		
$/person hour (VTTS)	Full	Rational	Difference	Full APS	Rational APS	Difference
Free flow time	$10.16	$10.51	-3.34%	$9.08	$9.73	-6.67%
Slowed down time	$10.88	$11.61	-6.31%	$10.96	$12.17	-9.91%
Trip time variability	$2.00	$2.12	-5.48%	$1.76	$2.47	-28.61%

3.3 Pairwise alternative rationality test

A weaker test is to compare the pairs of alternatives, so that even where an alternative from the full set might be chosen when it has no best attributes, it can still have a better level on a pairwise comparison. If the pair includes the reference alternative, it may be that this contrast delivers an outcome that passes a pairwise rationality test on more occasions. To our surprise, of the 54 choice sets that failed the strict rationality test (listed in Appendix), not one respondent chose the reference alternative, with 46 of the 54 choosing the third alternative. On closer inspection, of the 54 choice sets that failed the full choice set rationality test, all but one satisfied the pairwise rationality test, with 20 of the chosen alternatives having the better level on 5 attributes, 17 on four attributes, 14 on three attributes, and two on two attributes. This suggests that if a three-way and/or a two-way assessment of alternatives are both candidate processing strategies, then only one respondent failed both rationality tests on only one choice set.

Could it be that just as some researchers suggest that there is a bias towards the reference alternative, there might be circumstances where the bias is reversed? For modelling, it may be appropriate to remove the reference alternative and treat their processing strategy as elimination by alternatives, allowing the reference alternative to be specified as ‘non-existent’. This is equivalent to non-preservation of an alternative in contrast to an attribute. Within this dataset, 23 respondents chose the reference alternative for all 16 choice tasks while a further 17 respondents chose the alternative for 15 out of 16 choice tasks. However, with 70 respondents never choosing the reference alternative, total avoidance of the reference alternative was much more common than total avoidance of the two hypothetical alternatives.

[3] All parameter estimates are statistically significant in all four models.

[4] Details are available on request from the authors.

If an alternative passes the pairwise comparison test, that is, it is better on at least one attribute than the alternative to which it is compared, we can state that it is not dominated by the other alternative. Expressed another way, the alternative in question is dominated by the other alternative if, for every attribute, the attribute level is equal or worse than the other alternative. While the pairwise rationality test applied above to those who failed the three-way rationality test found only one case of dominance, an examination of *all observations* uncovered a wider pattern of choice of a dominated alternative. Over the dataset, 46 observations were found to be dominated. These are documented in Table 6. The first two columns indicate which alternative *dominated* the chosen alternative, i.e., which alternative was equal or better on all attributes, but still not chosen. Of note is the high number for alternative three, where one plausible explanation is that respondents are not paying as close attention to the third alternative, and hence missing a superior alternative. This explanation is supported by the results from the base multinomial logit model (see Table 9 below) where an alternative-specific constant for the second alternative is positive and significant, implying that, *ceteris paribus*, the second alternative is preferred to the third, and hence an ordering bias is at play.

Table 6: Dominance in the full sample

Alternative that dominated the chosen alternative		Consistent choice of alternatives over all 16 choice tasks	
Reference	10	Always chose reference alternative	9
SC Alternative 2	7	Never chose reference alternative	10
SC Alternative 3	28	Other	27
Reference and SC Alternative 2	1	-	-
Total	**46**	**Total**	**46**

To be truly effective, the dominance check requires an unlabelled experiment, such that the only points of comparison between alternatives are the attributes. In this experiment, while the two alternative routes are unlabelled, the reference alternative represents their current route, and thus other factors might be influencing whether they choose the reference alternative or one of the remaining two alternatives. For nine dominated observations, the respondent always chose the reference alternative over 16 choice tasks. This suggests that they were not trading over the attributes, such that a new alternative with superior attributes was not preferred. Conversely, for 10 dominated observations, the respondent never chose the reference alternative, instead trading only between the two hypothetical alternatives. In all 10 observations, the reference alternative dominated the chosen alternative. The respondent might have been dissuaded from the reference alternative by their actual experiences of it. Alternatively, inferences might be made about omitted attributes, leading to seemingly irrational choices being made (Lancsar and Louviere 2006). The remaining observations were by respondents who chose the reference alternative and a hypothetical alternative at least once each. We have no clear explanation for their choice of a dominated alternative. A preference for, or aversion to, the reference alternative might still have been in effect, except with some trading across these alternatives. Alternatively, the dominance might be the consequence of not paying attention, for example to the third alternative, as discussed above.

3.4 Influences of dominance and non-trading

It is often suggested that respondents are non-traders as a result of always selecting the same alternative, especially the reference alternative, across all choice sets. There are many reasons posited including lack of interest in the choice experiment, regret avoidance, and inertia. We investigated design attribute levels and respondent-specific characteristics as possible sources of influence in Table 7 (Model 1), where the binary dependent variable equals 1 for 23 observations who always choose the reference alternative, and zero otherwise for the remaining 355 respondents. Increased trip length decreases the probability of the respondent always choosing the reference alternative, as does a business trip purpose (in contrast to commuting and non-commuting). Two attributes that we have expected would be significant were not, namely the variability in total time as a percentage of the worst time for the reference alternative, and the percentage of total trip time in slowed down conditions.

We then ran a binary logit model (Model 2) to investigate possible systematic sources of influence on the choice of the reference alternative at a choice set level. This model delivered some very significant sources of influence, suggesting variety seeking behaviour (i.e., moving away from always choosing the reference alternative) as income increases, trip length increases, the trip is for business, the amount of toll road experience increases, and as there is engagement in attribute processing leading to an increasing number of attributes not being preserved. This latter evidence might be due the presence of greater engagement in evaluating the new alternatives. Also, with greater variability in travel times across the reference alternative, respondents are less likely to stay with the reference alternative, as expected. However, the sign for the percentage of time being in slowed down conditions is positive, which is the opposite effect to total time variability. This might suggest that there is relatively more congestion with shorter trips, which increases the probability of choosing the reference alternative.

Having identified some statistically significant influences on bias in favour of, or against, the reference alternative across all choice sets, and at a choice set level, we included the findings in the base models under full attribute relevance (Model 3) and under attribute non-preservation (Model 4). The overall log-likelihood for Model 3 improves from -5428 to -5331. The extra reference-alternative-specific characteristics were highly significant, and the reference constant became marginally significant and positive, suggesting that we have accounted for a growing number of the reasons why respondents do not chose the reference alternative. Similar improvements can be found for Model 4, with the log-likelihood improving from -5265 to -5173.

Table 7: Respondent and design influences on the choice of the reference alternative

	Full Relevance			**Ignored Attributes**
	Model 1	Model 2	Model 3	Model 4
	Reference Alternative chosen for all tasks	Reference Alternative chosen for single task	Base model with extra influences	Base model with extra influences
Constant	-1.4881 (-1.59)	1.1683 (9.34)	-	-
Time to complete a choice set (seconds)	-	0.0095 (7.95)	-	-
Trip length (kilometres)	-0.0293 (-2.34)	-0.0185 (-15.9)	-0.0107 (-8.70)	-0.0111 (-8.94)
Personal gross income ($'000s)	0.0102 (1.24)	-0.0034 (-3.21)	-0.0044 (-4.01)	-0.0042 (-3.72)
Business trip (compared to commuting and non-commuting)	-1.670 (-2.22)	-0.4048 (-6.78)	-0.3999 (-6.47)	-0.3995 (-6.34)
Ref alt time variability as percentage of Ref alt worst time	-1.6012 (-1.02)	-0.9469 (-4.86)	-1.1422 (-5.58)	-1.0013 (-4.84)
Percentage of total trip time in slowed down conditions	0.5060 (0.46)	0.3588 (2.46)	0.6835 (4.31)	0.4521 (2.92)
Amount of recent experience on toll roads (0-6)	-0.0147 (-0.11)	-0.0342 (-2.03)	-0.0465 (-2.65)	-0.0416 (-2.33)
Number of ignored attributes	0.1862 (0.94)	-0.0747 (-2.79)	-	-
Reference constant (1,0)	-	-	1.1828 (9.61)	1.1299 (9.11)
SC1 constant (1,0)	-	-	0.0730 (1.83)	0.0677 (1.69)
Free flow time (mins)	-	-	-0.0850 (-26.6)	-0.0904 (-26.65)
Slowed down time (mins)	-	-	-0.0953 (-15.3)	-0.1081 (-15.6)
Trip time variability (plus/minus mins)	-	-	-0.0067 (-1.14)	-0.0102 (-1.48)
Running cost ($)	-	-	-0.3906 (-20.7)	-0.4481 (-20.9)
Toll cost ($)	-	-	-0.5448 (-27.4)	-0.6303 (-30.7)
BIC	0.5357	1.3027	1.7817	1.7296
Log-likelihood at convergence	-77.50	-3930.20	-5331.12	-5173.80
Sample Size	378	6048	6048	6048

3.5 Dimensional vs holistic processing strategies

The pairwise test is reflective of the 'majority of confirming dimensions' rule (MCD) (Russo and Dosher 1983), which is concerned with the total count of superior attributes in each alternative. Pairs of attributes are compared in turn, with an alternative winning if it has a greater number of better attributes. The paired test continues until there is an overall winner. In our case, additionally, it might be that the reference alternative is dropped first, resulting in only a one pair test.

To test for the MCD heuristic in this dataset, a total count of best attributes was generated for each alternative, and then entered into the utility expressions for all three alternatives. To contribute to the count for an alternative, an attribute had to be *strictly better* than that attribute in all other alternatives in the choice set. That is, no ties were allowed[5]. The distribution of number of best attributes is shown in Table 8, both for the full relevance sample, and accounting for self-explication attribute non-preservation, with separate reporting for all alternatives and the chosen alternative only. The distribution for the chosen alternatives is skewed towards a higher number of best attributes in both cases, and higher means can also be observed. This alone does not suggest that MCD is being employed, as it would be expected that alternatives with a higher number of best attributes would also tend to have higher utilities.

Table 8: Number of strictly best attributes per alternative

	Full relevance				APS ignore			
	All alternatives		Chosen alternative		All alternatives		Chosen alternative	
Number of best attributes	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
0	2758	15.20%	467	7.72%	4703	25.92%	871	14.40%
1	8245	45.44%	2563	42.38%	8697	47.93%	2950	48.78%
2	5482	30.21%	2118	35.02%	3862	21.29%	1707	28.22%
3	1382	7.62%	709	11.72%	777	4.28%	439	7.26%
4	277	1.53%	191	3.16%	105	0.58%	81	1.34%
5	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	18144	100%	6048	100%	18144	100%	6048	100%
Mean	1.35		1.60		1.06		1.32	

Separate models were estimated allowing for ties, with similar but less significant results obtained. The model results are reported in Table 9, with the first column representing the base model, with all attributes assumed to be considered. The second column extends this base model, such that both the attribute levels and the number of best attributes impact on representative utility. The latter is highly significant, and positive in sign, so that as the number of best attributes increases, an alternative is more likely to be chosen, as would be expected. Additionally, an improvement in log-likelihood and Bayes Information Criterion (BIC) can be observed. The third column reports a model where only the number of best attributes and the alternative-specific constants are included, and the attribute levels are omitted. While the number of best attributes is highly significant, the model fit is considerably worse, suggesting that the number of best attributes cannot substitute for the attribute levels themselves.

[5] Accounting for ties did not materially affect the findings.

Table 9: Influence of majority of confirming dimensions

	Full Relevance			APS Ignore		
Reference constant (1,0)	0.0065 (0.13)	-0.0418 (-0.84)	0.5228 (15.96)	-0.0417 (-0.89)	-0.0797 (-1.67)	0.5149 (15.6)
SC1 constant (1,0)	0 .0749 (1.88)	0.0862 (2.16)	0.1339 (3.75)	0.0669 (1.67)	0.0821 (2.04)	0.1422 (3.95)
Free flow time (mins)	-0.0899 (-28.3)	-0.0853 (-26.0)		-0.0949 (-28.0)	-0.0884 (-24.9)	-
Slowed down time (mins)	-0.0963 (-16.1)	-0.0826 (-12.7)	-	-0.1146 (-16.9)	-0.0983 (-13.4)	-
Trip time variability (plus/minus mins)	-0.0177 (-3.07)	-0.0053 (-0.85)	-	-0.0184 (-2.68)	-0.0041 (-0.56)	-
Running cost ($)	-0.4147 (-22.2)	-0.3871 (-20.1)	-	-0.4735 (-22.4)	-0.4354 (-19.7)	-
Toll cost ($)	-0.5312 (-27.5)	-0.5274 (-27.4)	-	-0.6271 (-31.0)	-0.6123 (-30.2)	-
# of attributes in an alternative that are best	-	0.1041 (4.95)	0.3136 (19.79)	-	0.1269 (5.24)	0.4370 (23.9)
BIC	1.8051	1.803	2.0628	1.7514	1.7483	2.0295
Log-likelihood at convergence	-5428.17	-5417.55	-6224.89	-5265.81	-5252.05	-6123.98
Sample size	6048					
APS Ignored:						
Free flow time (mins)	-			944		
Slowed down time (mins)	-			1504		
Trip time variability (plus/minus mins)	-			2240		
Running cost ($)	-			1120		
Toll cost ($)	-			656		

While the second model is an improvement, its underlying form suggests that all respondents simultaneously consider and trade between both the attribute levels in a typical compensatory fashion, and the number of best attributes in each alternative. More plausibly, a respondent might resort solely to the MCD heuristic, or refrain from using it entirely. In recognition that there may be two classes of respondent, with heuristic application distinguishing between them, a latent class model[6] was estimated (Table 10). Two classes are defined[7], where the utility expressions in each class are constrained to represent one of the two heuristics. The first class contains the attribute levels and alternative-specific constants, as per the base model, while the second class contains only the number of best attributes. A further improvement in model fit is obtained with this model, with the log-likelihood improving from -5428.17 for the base model, to -5417.55 for the single class model that contains both the levels and the number of best attributes, to -5402.47 for the latent class model. Again the number of best attributes parameter is statistically significant and of the expected sign.

The same tests were performed, after accounting for the stated attribute non-preservation patterns of the respondents. Any ignored attributes were not included in the count of the number of best attributes. The fourth column of Table 9 sets out the base model that accounts for attribute non-preservation (or ignoring), which itself fits the data better than when all attributes are assumed to be attended to. The fifth column presents the model that accounts for both heuristics. The log-likelihood is smaller, at -5252.05 compared to -5265.81 for the base model, with the number of best attributes parameter being statistically significant and of the expected sign. The latent class model performs considerably better again, with a log-likelihood of -5218.52.

[6] See Hensher and Greene (2009) for other examples of the identification of attribute processing heuristics with the latent class model.

[7] We investigated a three-class model in which the additional class was defined by all attributes plus the number of best attributes. The overall fit of the model did not improve and many of the attributes were not statistically significant. We also estimated a three-class model with class-specific parameter estimates for attributes included in more than one class, but many parameters were not statistically significant. A further model allowing for random parameters was investigated but did not improve on the two-class model reported in Table 7.

Table 10: Identifying role of MCD - latent class model

	Full Relevance	**APS Ignore**
Class	1	1
Reference constant (1,0)	-0.4207 (-0.67)	-0.0676 (-1.06)
SC1 constant (1,0)	0.0674 (1.27)	0.0852 (1.51)
Free flow time (mins)	-0.1234 (-16.52)	-0.1448 (-16.6)
Slowed down time (mins)	-0.1192 (-11.37)	-0.1676 (-12.1)
Trip time variability (plus/minus mins)	-0.0145 (-1.83)	-0.0116 (-1.18)
Running cost ($)	-0.5467 (-15.04)	-0.6980 (-14.9)
Toll cost ($)	-0.7159 (-12.92)	-0.9038 (-18.0)
Class	2	2
# of attributes in an alternative that are best	0.2856 (2.76)	0.2665 (3.06)
Probability of class membership:		
Class 1	0.8465 (6.25)	0.8206 (9.58)
Class 2	0.1535 (6.35)	0.1794 (8.17)
BIC	1.7795	1.7287
Log-likelihood at convergence	-5402.47	-5218.52
Sample size	6048	
APS Ignored:		
Free flow time (mins)	-	944
Slowed down time (mins)	-	1504
Trip time variability (plus/minus mins)	-	2240
Running cost ($)	-	1120
Toll cost ($)	-	656

These results suggest that some respondents are employing the MCD heuristic. Under the heuristic, trading is not occurring on the absolute attribute levels. What matters instead is which alternative has the *best* level for each attribute, where tallies of the number of best attributes appear to act as a supplementary step when determining the best alternative. Overall, the mean probability of class membership of each class in both models is over 0.8 for processing of the constituent attributes and between 0.15 and 0.18 for the number of attributes being the determining influence.

The implication is that the application of the choice model must recognise that the trading amongst the attributes occurs up to a probability of 0.85 (or 0.82) on average, with the number of best attribute levels having an influence up to a probability of 0.15 (or 0.18) on average. This is an important finding that downplays the contribution of the marginal disutility of each attribute in the presence of the overall number of preferred attribute levels associated with an alternative.

3.6 Influence of the relative attribute levels

Another test relates to the relationship between the level of an attribute associated with the reference alternative and each of the other alternatives (Ref-SC1, Ref-SC2). We distinguished between differences where the reference alternative attribute level was better, equal and worse relative to SC1 and SC2. The choice response variable refers to the alternative chosen. A simple logit model was specified in which we included the best and worse attribute forms for all five design attributes (eliminating 'worst' for toll cost since there were no observations). The model is summarised in Table 11. Interpreting the parameter estimates is tricky. Where an attribute refers to a better level for the reference alternative (the difference for all attributes being negative on the attribute difference), a positive parameter estimate suggests that when the difference narrows towards zero, making the reference alternative relatively less attractive on that attribute, the probability of choosing a non-reference alternative (SC1 or SC2) increases. The parameter estimate is positive for 'better' except for trip time variability, producing the opposite behavioural response, which seems counter intuitive. The opposite behavioural response is found when the reference alternative is worse; all parameter estimates are positive suggesting that when the reference alternative becomes relatively less attractive (given it is worse), the probability of choosing SC1 or SC2 increases.

Table 11: Influence of referencing on choice response

6048 observations

Reference minus SC1 or SC2	Percent of data	Parameter estimates
Free flow time better	37.7	0.0915 (12.1)
Free flow time worse	62.3	0.0647 (7.45)
Slowed down time better	47.8	0.0860 (5.25)
Slowed down time worse	52.2	0.0770 (10.9)
Variability in time better	40.5	-0.0347 (-1.89)
Variability in time worse	59.5	0.0215 (1.84)
Running cost better	38.8	0.3090 (4.72)
Running cost worse	61.2	0.4996 (9.69)
Toll cost better	100	0.6336 (30.4)
Toll cost worse	0	-
Stated Choice Alternative 2 dummy (1,2)	-	0.1186 (2.96)
Log-likelihood at convergence	-3118.56	

3.7 Revision of the reference alternative

DeShazo (2002) suggested the idea of *reference point revision* in which preferences may be well-formed, but respondents' value functions shift when a non-status-quo option is chosen (see also McNair *et al.* 2010b). The shift occurs because the selection of a non-status-quo option is viewed as a transaction up to a probability, and this causes a revision of the reference point around which the asymmetric value function predicted by prospect theory is centred (Kahneman and Tversky, 1979)[8].

We ran a model in which we identified the chosen alternative from a previous choice set, and created a dummy variable equal to 1 if the chosen alternative in the previous choice set was a non-reference alternative. We then introduced, into the utility expressions, the lagged response variable indicating what the previous choice was, as a way of investigating the role of value-learning. We found (see Table 12) a mean estimate of 0.9357 (*t*-ratio of 15.73), which suggests that when a non-reference alternative is chosen, it revises the reference alternative in the next choice scenario, which increases the utility of the new 'reference' alternative. This is an important finding, supporting the hypothesis of DeShazo; it is also recognition of sequential interdependence between adjacent choice scenarios, which should be treated explicitly rather than through a correlated error variance specification, where the latter captures many unobserved effects at the alternative level.

Table 12: Identifying role of reference revision

Note: Choice set 1 is removed

	Full Relevance
Revised Reference (1,0)	*0.9358 (15.73)*
Free flow time (mins)	-0.01033 (-52.3)
Slowed down time (mins)	-0.0972 (-17.4)
Trip time variability (plus/minus mins)	-0.0178 (-2.96)
Running cost ($)	-0.4810 (-36.8)
Toll cost ($)	-0.6163 (-43.2)
BIC	1.7637
Log-likelihood at convergence	-5027.00
Sample size	5730

[8] Swait and Adamowicz (2001) also discuss MCD, but only seem to use it as an inspiration for the calculation of an entropy measure that represents choice task complexity

Conclusions

What does this evidence suggest for moving forward in the use of choice experiment data? We have identified a number of features of the choosing process that are associated with the design of the choice experiment, and the characteristics of respondents, that influence the stated choice outcome. Some very specific heuristics appear to have some systematic influence on choice, in particular the number of attributes that offer the best levels for an alternative, and the revision of the reference alternative as a result of value learning[9], reflected in a previous choice in the choice set sequence. Building both of these features into the estimated choice model seems to be a useful step forward in recognition of process rule heterogeneity. We also believe that the simple rationality test proposed herein for the entire choice set, and for pairwise alternatives, is a useful tool in eliminating data, if required, that has individuals choosing an alternative that has no single attribute that is better.

Another avenue for reconciling seemingly irrational choice behaviour stems from the recognition that the choice might be rational when a decision or process rule is employed by the decision maker. We have handled several decision rules in our analysis, namely the treatment of attributes as ignored using respondent self-explication, the application of the MCD heuristic, and revision of the reference alternative. However, other processes might be employed by the respondents that are not consistent with utility maximisation. For example, Gilbride and Allenby (2004) estimated a choice model that handled conjunctive and disjunctive screening rules, with choice treated as a compensatory process on the remaining alternatives. Here, a choice task that appears irrational might pass the rationality test after some alternatives have been eliminated in the screening stage. Swait (2009) allowed the unobserved utility of the choice alternatives to be in one of several discrete states. One of the states allowed conventional utility maximisation, while other states led to alternative rejection and alternative dominance. Again, rationality might prevail once the process rule is employed: in this case once rejection and dominance has been taken into account. We propose that one way to assess these and other new model forms is to determine how well they can explain decisions that appear irrational when viewed through the conventional prism of utility maximisation.

Of interest to the analyst are possible ways in which irrational behaviour can be minimised in a stated choice environment. In our data, there appeared to be no link between the task order number and the rate of irrational behaviour, which suggests that the number of choice tasks might not have an impact, within reasonable limits. Choice task complexity (number of alternatives, attributes and attribute levels) was not varied in this analysis; however the impact of task complexity on irrational behaviour would be an interesting area of research. Also of interest is the rationality of choice in market conditions, which may be impacted by habit, mood, time pressure, and ease with which information can be compared. We anticipate that these influences would lead to a decrease in rationality of choice, either through an increase in errors, or an increase in use of decision rules and heuristics. If the aim of a stated choice task is to successfully predict market choices, encouraging rational choice in the stated choice environment might not actually be the best way forward. Survey realism might instead be more important.

We present a model below as a contrast to the base model (Table 9, column 2 and column 5), where we include reference revision, majority of conforming dimensions, and eliminate choice

[9] Value learning in its broadest meaning implies underlying preferences are changing. Reference revision, on the other hand, can occur when preferences are stable but the objective is to maximise the likelihood of implementation of the most preferred alternative observed *over the course of the sequence of questions*. In a sense the latter is a special case of the former. We thank Ben McNair for pointing this out.

sets that fail the three-way and two-way rationality tests[10]. Accommodating value learning through reference revision involves treating the first choice set differently; to allow for this we, introduce a dummy variable for the initial reference alternative for choice one only. We also include design and contextual variables that are correlates, to some degree, with the presence of non-trading in terms of always selecting the existing (i.e., non-revised) reference alternative across all 16 choice sets, or selection of the existing reference alternative in a specific choice set.

Table 13: Revised full model for applications

	Ignored Attributes
Trip length (kilometres)	-0.0098 (-7.54)
Personal gross income ($'000s)	-0.0077 (-7.46)
Business trip (compared to commuting and non-commuting)	-0.3490 (-5.27)
Existing reference alternative time variability as percentage of worst time	-0.8548 (-3.91)
Percentage of total trip time in slowed down conditions	0.5703 (3.40)
Amount of recent experience on toll roads (0-6)	-0.0304 (-1.61)
Free flow time (mins)	-0.0909 (-23.6)
Slowed down time (mins)	-0.0938 (-12.04)
Trip time variability (plus/minus mins)	0.0103 (1.34)
Running cost ($)	-0.4539 (-19.0)
Toll cost ($)	-0.6414 (-29.4)
# of attributes in an alternative that are best	**0.2646 (10.0)**
Reference revision (1,0)	**0.8843 (13.8)**
Initial Choice Set Reference dummy (1,0) for choice sets 2-16	**1.1442 (8.99)**
BIC	1.6092
Log-likelihood at convergence	-4600.45
Sample Size	5793

The mean estimates of value of travel time savings in Table 13 for free flow time are $12.02 based on the running cost parameter, and $8.50 based on the toll cost parameter. The equivalent VTTS for slowed down time are $12.40 and $8.77 per person hour. These estimates can be contrasted with the findings of the 'base' model (reported in Table 5) which only included the design attributes and constants for the existing reference alternative (without value learning), namely $11.62, $9.73 for free flow time, and $14.53 and $12.17 for slowed down time. In all but the valuation for free flow time with respect to running cost, when the additional influences in Table 13 are not taken into account, we find non-marginal over-estimation of the mean VTTS.

This paper will hopefully engender an interest in further inquiry into the underlying sources of process heterogeneity that should be captured explicitly in the formulation of the utility expressions that represent the preference domain of each respondent for each alternative. Including additional attribute and alternative-processing related explanatory variables will provide plausible explanations of utility maximising behaviour in choice making.

[10] In this particular data set, eliminating choice sets that fail the three-way and two-way rationality tests had so significant influence at all on the parameter estimates compared with including these data points.

Appendix

Alternative	FF	SDT	VAR	RC	TC	TotT	TotC	choice1
1 (SQ)	20	0	0	2.6	0	20	2.6	0
2	**17**	2	6	2.21	4	19	6.21	0
3	26	3	4	2.99	1	29	3.99	1
1 (SQ)	8	2	2	1.2	0	10	1.2	0
2	7	1	6	1.02	4	8	5.02	0
3	10	2	4	1.38	1	12	2.38	1
1 (SQ)	60	0	30	7.8	0	60	7.8	0
2	51	2	34	6.63	0.5	53	7.13	0
3	69	2	34	8.97	3	71	11.97	1
1 (SQ)	25	0	18	3.25	0	25	3.25	0
2	29	4	12	2.28	0.5	33	2.78	0
3	29	3	15	4.23	3	32	7.23	1
1 (SQ)	30	0	5	3.9	0	30	3.9	0
2	39	2	6	4.49	0.5	41	4.99	1
3	34	4	6	3.32	2.5	38	5.82	0
1 (SQ)	22	0	2	2.86	0	22	2.86	0
2	29	2	6	3.29	0.5	31	3.79	1
3	25	4	6	2.43	2.5	29	4.93	0
1 (SQ)	16	0	5	2.08	0	16	2.08	0
2	21	2	6	2.39	0.5	23	2.89	1
3	18	4	6	1.77	2.5	22	4.27	0
1 (SQ)	20	0	5	2.6	0	20	2.6	0
2	26	2	6	2.99	0.5	28	3.49	1
3	23	4	6	2.21	2.5	27	4.71	0
1 (SQ)	22	0	2	2.86	0	22	2.86	0
2	29	3	4	3.29	1	32	4.29	1
3	15	2	4	3.72	3.5	17	7.22	0
1 (SQ)	35	10	2	5.33	0	45	5.33	0
2	46	8	4	6.13	1	54	7.13	1
3	24	7	4	6.93	3.5	31	10.43	0
1 (SQ)	8	2	2	1.2	0	10	1.2	0
2	10	2	4	1.38	1	12	2.38	1
3	6	1	4	1.55	3.5	7	5.05	0
1 (SQ)	40	5	8	5.59	0	45	5.59	0
2	28	6	5	7.27	3	34	10.27	0
3	34	6	6	7.27	0.5	40	7.77	1
1 (SQ)	50	10	10	7.28	0	60	7.28	0
2	35	13	7	9.46	3	48	12.46	0
3	42	13	8	9.46	0.5	55	9.96	1
1 (SQ)	25	5	8	3.64	0	30	3.64	0
2	18	6	5	4.73	3	24	7.73	0
3	21	6	6	4.73	0.5	27	5.23	1
1 (SQ)	45	45	22	9.36	0	90	9.36	0
2	32	58	16	12.17	3	90	15.17	0
3	38	58	19	12.17	0.5	96	12.67	1
1 (SQ)	40	40	35	8.32	0	80	8.32	0
2	28	52	24	10.82	3	80	13.82	0
3	34	52	30	10.82	0.5	86	11.32	1
1 (SQ)	15	0	2	1.95	0	15	1.95	0
2	10	4	4	2.54	3	14	5.54	0
3	13	4	4	2.54	0.5	17	3.04	1
1 (SQ)	35	40	12	7.67	0	75	7.67	0
2	24	52	9	9.97	3	76	12.97	0
3	30	52	11	9.97	0.5	82	10.47	1
1 (SQ)	10	25	12	3.25	0	35	3.25	0
2	7	32	9	4.23	3	39	7.23	0
3	8	32	11	4.23	0.5	40	4.73	1
1 (SQ)	22	0	2	2.86	0	22	2.86	0
2	15	4	4	3.72	3	19	6.72	0
3	19	4	4	3.72	0.5	23	4.22	1
1 (SQ)	30	7	6	4.45	0	37	4.45	0
2	21	9	5	5.78	3	30	8.78	0
3	26	9	6	5.78	0.5	35	6.28	1
1 (SQ)	90	0	45	11.7	0	90	11.7	0
2	63	4	32	15.21	3	67	18.21	0
3	76	4	38	15.21	0.5	80	15.71	1
1 (SQ)	65	25	15	10.4	0	90	10.4	0
2	46	32	10	13.52	3	78	16.52	0
3	55	32	13	13.52	0.5	87	14.02	1
1 (SQ)	55	5	12	7.54	0	60	7.54	0
2	38	6	9	9.8	3	44	12.8	0
3	47	6	11	9.8	0.5	53	10.3	1
1 (SQ)	20	20	10	4.16	0	40	4.16	0

Alternative	FF	SDT	VAR	RC	TC	TotT	TotC	choice1
1 (SQ)	25	0	18	3.25	0	25	3.25	0
2	18	4	12	4.23	3	22	7.23	0
3	21	4	15	4.23	0.5	25	4.73	1
1 (SQ)	55	10	15	7.93	0	65	7.93	0
2	38	13	10	10.31	3	51	13.31	0
3	47	13	13	10.31	0.5	60	10.81	1
1 (SQ)	240	30	30	33.54	0	270	33.54	0
2	168	39	21	43.6	3	207	46.6	0
3	204	39	26	43.6	0.5	243	44.1	1
1 (SQ)	30	15	8	5.07	0	45	5.07	0
2	21	20	5	6.59	3	41	9.59	0
3	26	20	6	6.59	0.5	46	7.09	1
1 (SQ)	30	15	40	5.07	0	45	5.07	0
2	21	20	28	6.59	3	41	9.59	0
3	26	20	34	6.59	0.5	46	7.09	1
1 (SQ)	35	10	2	5.33	0	45	5.33	0
2	24	13	4	6.93	3	37	9.93	0
3	30	13	4	6.93	0.5	43	7.43	1
1 (SQ)	15	5	6	2.34	0	20	2.34	0
2	10	6	5	3.04	3	16	6.04	0
3	13	6	6	3.04	0.5	19	3.54	1
1 (SQ)	15	5	2	2.34	0	20	2.34	0
2	10	6	4	3.04	3	16	6.04	0
3	13	6	4	3.04	0.5	19	3.54	1
1 (SQ)	25	5	18	3.64	0	30	3.64	0
2	18	6	12	4.73	3	24	7.73	0
3	21	6	15	4.73	0.5	27	5.23	1
1 (SQ)	40	5	8	5.59	0	45	5.59	0
2	28	6	5	7.27	3	34	10.27	0
3	34	6	6	7.27	0.5	40	7.77	1
1 (SQ)	20	10	8	3.38	0	30	3.38	0
2	14	13	5	4.39	3	27	7.39	0
3	17	13	6	4.39	0.5	30	4.89	1
1 (SQ)	25	10	10	4.03	0	35	4.03	0
2	18	13	7	5.24	3	31	8.24	0
3	21	13	8	5.24	0.5	34	5.74	1
1 (SQ)	25	5	8	3.64	0	30	3.64	0
2	18	6	5	4.73	3	24	7.73	0
3	21	6	6	4.73	0.5	27	5.23	1
1 (SQ)	17	3	4	2.44	0	20	2.44	0
2	12	4	2	3.18	3	16	6.18	0
3	14	4	3	3.18	0.5	18	3.68	1
1 (SQ)	45	15	15	7.02	0	60	7.02	0
2	32	20	10	9.13	3	52	12.13	0
3	38	20	13	9.13	0.5	58	9.63	1
1 (SQ)	30	10	10	4.68	0	40	4.68	0
2	21	13	7	6.08	3	34	9.08	0
3	26	13	8	6.08	0.5	39	6.58	1
1 (SQ)	35	10	8	5.33	0	45	5.33	0
2	24	13	5	6.93	3	37	9.93	0
3	30	13	6	6.93	0.5	43	7.43	1
1 (SQ)	8	2	2	1.2	0	10	1.2	0
2	6	3	4	1.55	3	9	4.55	0
3	7	3	4	1.55	0.5	10	2.05	1
1 (SQ)	17	3	8	2.44	0	20	2.44	0
2	12	4	5	3.18	3	16	6.18	0
3	14	4	6	3.18	0.5	18	3.68	1
1 (SQ)	20	5	12	2.99	0	25	2.99	0
2	14	6	9	3.89	3	20	6.89	0
3	17	6	11	3.89	0.5	23	4.39	1
1 (SQ)	50	40	15	9.62	0	90	9.62	0
2	35	52	10	12.51	3	87	15.51	0
3	42	52	13	12.51	0.5	94	13.01	1
1 (SQ)	22	3	4	3.09	0	25	3.09	0
2	15	4	2	4.02	3	19	7.02	0
3	19	4	3	4.02	0.5	23	4.52	1
1 (SQ)	20	10	15	3.38	0	30	3.38	0
2	14	13	10	4.39	3	27	7.39	0
3	17	13	13	4.39	0.5	30	4.89	1
1 (SQ)	90	0	15	11.7	0	90	11.7	0
2	63	4	10	15.21	3	67	18.21	0
3	76	4	13	15.21	0.5	80	15.71	1
1 (SQ)	50	10	15	7.28	0	60	7.28	0

References

Bateman I.J., Carson R.T., Day B., Dupont D., Louviere J.J., Morimoto S., Scarpa R. et al, (2008). Choice set awareness and ordering effects in discrete choice experiments. CSERGE Working Paper EDM 08-01.

Cameron, T.A. and J.R. DeShazo (2008) Differential Attention to Attributes in Utility-theoretic Choice Models (mimeo)

Chorus, C.G. (2010) Random Regret Minimization: An improved model form and comparisons with the Multinomial Logit model, Delft University of Technology, The Netherlands.

Day B. and Pinto J.L. (2010) Ordering anomalies in choice experiments. *Journal of Environmental Economics and Management* doi:10.1016/j.jeem.2010.03.001.

Day B, Bateman I J, Carson R T, Dupont D, Louviere J J, Morimoto S, Scarpa R et al (2009) Task Independence in Stated Preference Studies: A Test of Order Effect Explanations. CSERGE Working Paper EDM 09-14,

DeShazo J.R. (2002) Designing transactions without framing effects in iterative question formats. *Journal of Environmental Economics and Management* 43, 360-385.

Fiebig, D., Keane, M., Louviere, J., and Wasi, N. (2009) The generalized multinomial logit: accounting for scale and coefficient heterogeneity, *Marketing Science*, published online before print July 23, DOI:10.1287/mksc.1090.0508

Gilbride, T. J. and Allenby, G. M. (2004) A choice model with conjunctive, disjunctive, and compensatory screening rules. Marketing Science, 23(3), 391-406.

Greene, W.H. and Hensher, D.A. (2010) Does scale heterogeneity matter? a comparative assessment of logit models, Transportation, 37 (3), 413-428.

Haaijer, R., Wagner, K., and Wedel M. (2000) Response Latencies in the Analysis of Conjoint Choice Experiments, Journal of Marketing Research, Vol. XXXVII (August), 376-382.

Hensher, D.A. (2010) Attribute processing, heuristics and preference construction in choice analysis, in Hess, S. and Daly, A. (eds.) *State-of Art and State-of Practice in Choice Modelling*, Emerald Press, UK., 35-70.

Hensher, D. A. and Greene, W. H. (2009) Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification. *Empirical Economics*, published online before print August 22, DOI: 10.1007/s00181-009-0310-x.

Hensher, D.A., Greene, W.H. and Chorus, C. (2010) Random regret minimisation or random utility maximisation: an exploratory analysis in the context of automobile choice, submitted to *Transportation Research E.*

Hess, S. and Hensher, D (in press 2010) Using conditioning on observed choices to retrieve individual-specific attribute processing strategies, *Transportation Research Part B.*

Kahneman D. and Tversky A. (1979) Prospect theory: An analysis of decisions under risk. *Econometrica* 47, 263-291.

Lancsar, E. and Louviere, J.J. (2006) Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences?, *Health Economics*, 15, 797-811.

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

McNair B J, Bennett J and Hensher D A (2010a) A comparison of responses to single and repeated discrete choice questions. *Occasional Paper #14*, Crawford School of Economics and Government, The Australian National University, Canberra.

McNair B J, Hensher D A and Bennett J (2010b) Modelling heterogeneity in response behaviour towards a sequence of discrete choice questions: a latent class approach. *Occasional Paper #16*, Crawford School of Economics and Government, The Australian National University, Canberra.

Racevskis L. and Lupi F. (2008) Incentive Compatibility in an Attribute-Based Referendum Model. Paper presented at American Agricultural Economics Association Annual Meeting, Orlando, FL, 27-29 July.

Rose, J. and Bliemer, M. (2008) Stated preference experimental design strategies, in D. A. Hensher and K. J. Button (eds). *Handbook of Transport Modelling* Oxford, Elsevier: 151-179.

Rose, J.M. and Black, I. (2006) Means Matter, but Variance Matter Too: Decomposing Response Latency Influences on Variance Heterogeneity in Stated Preference Experiments, Marketing Letters, 17(4), 295-310.

Rose, J., Bliemer, M., Hensher, D. A. and Collins, A. (2008) Designing efficient stated choice experiments in the presence of reference alternatives. Transportation Research Part B 42(4): 395-406.

Russo, J. E., and Dosher, B. A. (1983) Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 9(4), 676-696.

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

Swait, J. (2009) Choice models based on mixed discrete/continuous PDFs. *Transportation Research Part B*, 43, 766–783.