

Hensher, David A.; Collins, Andrew T.

Article

Interrogation of responses to stated choice experiments: Is there sense in what respondents tell us? : a closer look at what respondents choose and process heuristics used in stated choice experiments

Journal of Choice Modelling

Provided in Cooperation with:

Journal of Choice Modelling

Suggested Citation: Hensher, David A.; Collins, Andrew T. (2011) : Interrogation of responses to stated choice experiments: Is there sense in what respondents tell us? : a closer look at what respondents choose and process heuristics used in stated choice experiments, Journal of Choice Modelling, ISSN 1755-5345, University of Leeds, Institute for Transport Studies, Leeds, Vol. 4, Iss. 1, pp. 62-89

This Version is available at:

<https://hdl.handle.net/10419/66832>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc/2.0/uk/>

Interrogation of Responses to Stated Choice Experiments: Is there sense in what respondents tell us?

A Closer Look at what Respondents Choose and Process Heuristics used in Stated Choice Experiments

David A. Hensher^{1,*} Andrew T. Collins^{1,†}

¹Institute of Transport and Logistics Studies, The University of Sydney Business School,
The University of Sydney, NSW 2006 Australia

Received June 2010, revised versions received October 2010, accepted March 2011

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 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 meaningful reason. We find that there is relatively less effort placed on looking closely at the data defining each choice situation for each respondent, as constructed by the stated choice experiment, and seeing if there exist 'evidential' rules that support in a plausible way, the choice responses. Heuristics investigated in this paper that might aid in our understanding of how choice scenarios are processed, leading to a choice outcome, include the role of dimensional vs. holistic attribute processing, the influence of relative attribute levels, and the revision of the reference alternative as value learning across sequenced choice sets. 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 future inclusion in the estimation of all choice models. The evidence suggests that there is a great deal of behavioural sense in stated choice responses, for all manner of possible reason.

Keywords: Choice experiments, plausible choice, heuristics, sequence effects, referencing, reference revision, value learning, attribute processing, majority of confirming dimensions, evidential rules

1 Introduction

Stated choice experiments are used extensively to create data capable of modelling choices in order 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 for 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. 2010 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 2010, and Cameron and DeShazo 2010), using a variety of supplementary questions on how attributes are processed and/or developing model functional-forms that capture specific heuristics. 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) as well as 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 choice set level, and set up candidate rules, or heuristics (often referred to as ‘rules of thumb’) that align with one or more possible processing rules used by an individual, within and between sequentially administered choice sets, to reveal their choice response. Specifically, the analysis herein is looking for evidence that would be consistent with respondents’ use of heuristics to make choices in stated choice experiments. This matters because of the small, but accumulating empirical evidence, that alternative attribute processing strategies (APS) do influence behavioural outputs such as estimates of willingness to pay and model predictive capability (see Hensher 2010 for an overview). While we can never be certain that a specific 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 conducted¹, 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). We begin with the most commonly assumed normative processing rule which assumes (in the

¹ We undertook this exact same exercise on a number of data sets and a number of respondents in each data set and the message was the same or very similar.

absence of any known attribute processing heuristic) that all attributes (and levels) are relevant, and that a fully compensatory processing strategy is active at the choice set level. Focussing on these five attributes only, we highlight in shaded grey the most attractive attribute level (e.g., lowest FF), which varies across the attributes, and propose that if an alternative had the most attractive level on one or more attributes, and that alternative was chosen, then we can reasonably suggest that the respondent was 'plausible' in their choice, assuming that the heuristic being used to process the choice set preserves (i.e., does not ignore) the attribute(s) with the 'most attractive level(s)' based, of course, on only the offered attributes. Applying the same logic across all of the sixteen choices that each respondent made, we found that 51 of the 300 respondents are consistently selecting options which are best on the same attribute, where the experimental design does not allow them to consistently choose such that two or more attributes are always best.

There could be other reasons why an alternative is chosen, regardless of the attribute levels and their relative performance, 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). Indeed, if a respondent focuses on only one attribute, then we might be observing a consistent elimination-by-aspects heuristic². However, on the face of the observed attribute evidence, the 16 choice scenarios satisfy a 'plausible choice' test in 16 situations. Five of the choice scenarios show the status quo as the preferred alternative (bolded in the choice column). It may also be that this example individual adopts one or more attribute processing (AP) 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 'plausibility' test used above. We investigate a number of these AP rules in the following sections.

Furthermore, supplementary data associated with the respondents' perception of 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 responses to supplementary questions. Looking at the possibility that this individual may also have added up FF and SDT and/or Cost and Toll, we cannot find any evidence within the 'plausible choice' test that it would have failed if attribute addition (TotTime, TotCost) had not been applied, although this may have assisted in making the choice. The following sections of the paper undertake a more formal inquiry using another data set collected in 2007 in New Zealand, to delve more deeply into alternative 'plausible choice' tests as well as the role of non-compensatory heuristics in aiding our understanding of how stated choice sets are processed in assisting the selection of a choice outcome. 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 (or heuristics) that explain specific choice responses under specific assumptions. The investigated rules and tests focus on the influence of the choice sequence on choice response, a pairwise alternative plausibility test and the presence of dominance, the influence of non-trading, dimensional vs. holistic attribute processing, the influence of relative attribute levels, and revision of the reference alternative as value learning across sequenced choice sets. We then discuss the evidence, and conclude with a proposal to include two new explanatory variables in choice models to capture the number of attributes in an alternative that are 'best' as well as value learning, together

² We acknowledge a referee for this observation.

with a statement of the degree of confidence one might have in the behavioural sense of the data emanating from a stated choice experiment.

Table 1: An Example of 16 Choice Scenario Responses Evaluated by one Respondent

Choice scenario	Alternative	TotTime	TotCost	Var	FF	SDT	Cost	Toll	Choice	Plausible = 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

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 136 commuters, 116 non-commuters and 126 individuals travelling on employer business in the catchment area around Tauranga in the North Island of New Zealand. A stated choice experiment was included together with questions that sought information on a recent trip which was used to construct both the reference (i.e., status quo) alternative, and the two other alternatives, which had levels that pivoted around the status quo alternative. 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. The experimental design was composed of two blocks of 16 choice scenarios each, and can be found in full in Appendix A. Respondents were randomly assigned to one of the two blocks, with the order in which the 16 choice scenarios were presented also randomised. The levels of the design were optimised in accordance with efficient design theory, with a d-error measure employed (see Rose and Bliemer 2008 for details).

Table 2: Profile of the Attribute Range in the Choice Experiment

Attribute	Levels
Free Flow Time (variation around reference level)	-30%, -15%, 0, 15%, 30%
Slowed Down Time (variation around reference level)	-30%, -15%, 0, 15%, 30%
Trip Time Variability	±0%, ±5%, ±10%, ±15%
Running Cost (variation around reference level)	-40%, -10%, 0, 20%, 40%
Toll Cost (see note (ii) in text)	\$0, \$0.5, \$1, \$1.5, \$2, \$2.5, \$3, \$3.5, \$4

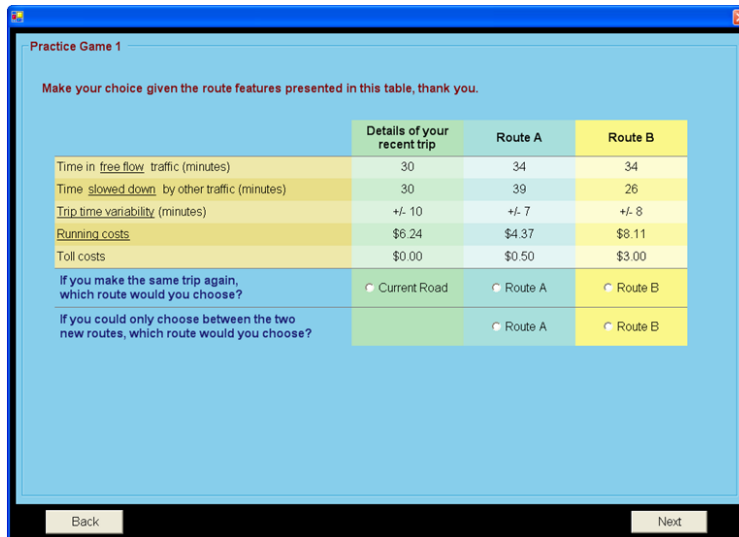


Figure 1: An example of a Stated Choice Screen

A few additional rules were imposed on the design:

- (i) Free flow and slowed times³ in the non-reference alternatives were set to a base of five minutes if the respondent entered zero for their current trip;
- (ii) To obtain sensible trip time variability levels in minutes, we asked respondents to suggest a range of departure times experienced to ensure they arrived at their destination at the planned arrival time. This range was used to identify the actual trip time variability given the percentages used in the design. Where the departure times were reported as the same as the recent reference trip, we set an artificial base as per the same rule in (i).
- (iii) Given that tolled routes were currently not available at the time of the survey, the proposed new tolled route was assigned a range of values. Construction of the new toll road was approved in 2010 with a proposed fixed toll in the range evaluated.

In addition, supplementary questions were asked upon completion of all 16 choice scenarios on whether specific attributes were ignored. The entire survey instrument was programmed as a Computer Assisted Personal Interview (CAPI), that enables the attribute levels to be tailored to (i.e., pivoted around) each respondent's recent trip experience. An interviewer was present and guided the respondents through the survey screens. All data is automatically captured in a database. The software has built in checks to ensure that all data provided were logical where appropriate (e.g., the travel time, given distance, delivered a meaningful average trip speed). Given the focus of this paper, other details of the study are not provided.

3 Investigating Candidate Evidential Rules

As a prelude to investigating a number of candidate heuristics (or evidential rules) that might contribute to explaining choice response, we continue the theme of 'plausible choice' in the contexts of full attribute relevance and omitting those attributes which the respondent claimed not to have considered. The following analyses are performed at both the choice set and respondent level, where we use the word 'observation' to refer to a choice set assessment, and 'respondent' to refer to the assessment over all (16) choice sets, with the latter providing evidence that those respondents who fail the various tests are exhibiting different behavioural tendencies overall, not just in response to a specific feature of the experimental design. We assess the implications of the evidence on willingness to pay (WTP) estimates, before investigating five speculative but interesting heuristics associated with (i) pairwise alternative plausibility and the presence of dominance, (ii) the influence of non-trading, (iii) the role of dimensional vs. holistic attribute processing, (iv) the influence of relative attribute levels, and (v) the revision of the reference alternative as value learning across sequenced choice sets.

The 'plausible choice' test presented above for one respondent can be applied across the 6,048 observations in the New Zealand data. Appendix B details all 54 choice sets (or scenarios) where the test failed. An alternative that would fail the test if chosen was present in 291 choice scenarios, resulting in a failure rate of 18.6 percent.

³ 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.

Note that the lack of a toll in some alternative (i.e., non-tolled routes) meant that the reference alternative always had at least one best attribute, and so if it was chosen, the ‘plausible choice’ test could not fail. Table 3 also shows the proportion (and counts) of plausible choice sets by choice task sequence number. When all attributes are assumed to be relevant, we find, across all 16 choice sets, that 99.12 percent of the observations pass the ‘plausible choice’ test associated with one or more attributes being best on the chosen alternative (with the percentage varying across the 16 choice sets from 100 percent to 98.4 percent). When we omit those attributes which the respondent claimed not to have considered, i.e., they were ignored, 95.78 percent of the observations pass this test (with the percentage varying across the 16 choice sets from 98.41 percent to 94.44 percent), suggesting that regardless of respondents’ claims of attributes being ignored or not, there is a very high incidence of plausible choosing. The evidence also suggests that there is no noticeable deterioration in plausible choice response as the respondent works through the choice sets from set 1 to set 16. At the respondent level, we find that the 54 choice observations that failed the ‘plausible choice’ test were spread across 49 respondents.

The structure of the design has an impact on the incidence of observations that fail the ‘plausible choice’ test. If full attribute attendance is assumed, then the test cannot be failed if every alternative in the experimental design has at least one best attribute. In this empirical setting, only one alternative in the design did not have a best level (choice scenario 31 in Appendix A), which might have had some role in keeping the incidence rate low (54 observations out of a possible 291). Other choice scenarios also allowed the test to fail as a consequence of the forced variability in slowed down and free flow time when the recent trip values were less than five minutes (rule (i) discussed earlier). Once ignored attributes are taken into account, the number of scenarios in which the test is failed can in no way be inferred from the experimental design. While there are a finite (albeit large) number of combinations with which the attributes can be ignored or preserved, the analyst does not know *a priori* which of these will be chosen. Looking to the entire dataset, it can be determined that when accounting for the reported ignoring of attributes, 255 observations are implausible out of a possible 1,699 choice scenarios where an implausible choice could have been made, spread across 99 respondents.

We also ran two simple logit models (not reported herein) to explore the possible influence of the commuter’s age, income and gender on whether the choice response for each choice set was plausible (1) or not (0) under the ‘plausible choice’ test. One model assumed full attribute relevance and the other accounted for the attributes that the respondents stated as ignored (or not preserved⁴). Income and gender had no influence, but age had a statistically significant impact when accounting for whether an attribute was ignored or not, with the probability of satisfying the ‘plausible choice’ test increasing as the commuter ages.

⁴ We are starting to see, in the literature, a number of ways of indicating that attributes are ignored. A popular language, especially in the environmental literature, is ‘attribute non-preservation’ or ‘attribute non-attendance’.

Table 3: Influence of Choice Sequence on Choice Response

Choice Set Sequence	Assuming full attribute relevance		Allowing for attribute being ignored	
	Proportion plausible	Count non-plausible	Proportion plausible	Count non-plausible
1	0.9894	4	0.9471	20
2	1.0000	0	0.9497	19
3	0.9894	4	0.9392	23
4	0.9894	4	0.9603	15
5	0.9974	1	0.9524	18
6	0.9841	6	0.9603	15
7	0.9921	3	0.9841	6
8	0.9947	2	0.9550	17
9	0.9841	6	0.9444	21
10	0.9894	4	0.9656	13
11	0.9894	4	0.9603	15
12	0.9921	3	0.9550	17
13	0.9894	4	0.9550	17
14	0.9894	4	0.9524	18
15	0.9894	4	0.9709	11
16	1.0000	0	0.9735	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 ‘plausible choice’ test (both under full attribute relevance and allowing for attributes to be ignored) and the amount of time to complete each of the 16 choice scenarios (i.e., the response latency). Statistically significant relationships were found between the choice scenario completion time and the ‘plausible choice’ test, both under full attribute relevance and when attribute processing was taken into account, and are reported in Table 4 (i) at the choice set level, and Table 4 (ii) at the respondent level. We find that for respondents who satisfied the ‘plausible choice’ test at the choice set level, the average time to complete a choice set was 27.47 seconds, with a standard deviation of 26.03 seconds; however when we account for the choice set response being implausible at the observation level, we find that the mean time decreases by 5.21 seconds under full attribute relevance and 5.58 seconds when ignoring attributes is accounted for. When we do the same comparison at the respondent level, we find for respondents who have at least one choice set not satisfying the plausibility test, that the average time to complete a choice screen decreases by 4.84 and 3.66 seconds for full attribute relevance and attribute non-attendance respectively, relative to the respondents who pass the test. One possible explanation for this difference in completion time is that those who pass the ‘plausible choice’ test are more engaged in the choice task. Alternatively, those who fail the test might be employing some other heuristic that allows them to make a more rapid choice. Clearly, no definitive causal inferences can be drawn, despite speculative opinion that such respondents might be less engaged in the task.

Table 4: Influences on Choice Scenario Completion Time
Simple OLS Regression

(i) Choice set level		
	Full attribute relevance	Allowing for attribute being ignored
Constant	22.2963 (17.1)	22.1163 (30.8)
Full Relevance Plausible Choice Test (1,0)	5.2159 (3.96)	
Plausible Choice Test under Attribute Non-Preservation (1,0)	-	5.5856 (7.5)
Adjusted R ²	0.00035	0.0019
Sample size	6,048	
(ii) Respondent level		
	Full attribute relevance	Allowing for attribute being ignored
Constant	23.2659 (53.9)	24.7287 (73.63)
Full Relevance Plausible Choice Test (1,0)	4.8474 (10.1)	
Plausible Choice Test under Attribute Non-Preservation (1,0)	-	3.6691 (9.01)
Adjusted R ²	0.00040	0.0037
Sample size	6,048	

3.1 Derivative Willingness to Pay

The next task is to estimate choice models at the choice set and respondent level that distinguish between (i) the full sample (6,048 observations or 378 respondents) assuming all attributes are relevant (Full), (ii) the full sample with choice scenarios removed when the 'plausible choice' test failed (5,994 observations, 329 respondents) (Plausible), (iii) the full sample taking into account attribute ignoring as an attribute processing strategy (6,048 observations, 378 respondents) (Full APS), and (iv) the full APS sample with choice scenarios removed when the 'plausible choice' test failed (5,793 observations, 279 respondents) (Plausible APS). The findings on values of travel time savings (VTTS) are summarised in Table 5⁵. We have also included the percentage changes in the mean VTTS estimates as a way of identifying the behavioural implications of failing the 'plausible choice' 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.

At the choice set level (Table 5(i)), while the differences are marked in some cases, none of the differences in mean VTTS are statistically different, especially the weighted average VTTS (where the weights relate to the attribute levels for free flow and slowed down time, and running and toll cost), using the delta test to obtain standard errors⁶. This is the case even when over four percent of the sample is removed due to a suspicion of implausible choice behaviour. This finding suggests that the underlying model is robust, and able to cope with a small percentage of seemingly implausible decisions. However, when we compare the mean VTTS at the respondent level (in Table 5(ii)), we find statistically significant differences, given standard errors calculated using the delta method and 1,000 random draws. This is an important finding, suggesting that the behavioural implications in terms of VTTS is not of concern when we focus on individual choice sets, but when we remove entire respondents who fail the plausible test on at least one choice set, the differences are

⁵ All parameter estimates are statistically significant in all four models.

⁶ Details are available on request from the authors.

significant. The respondent-level evidence supports the findings of Scarpa et al. (2007) who found that the WTP estimates are of a different magnitude when ‘irrational’ respondents were removed, which were of a considerably higher proportion than in the current study.

Table 5: Implications of the ‘Plausible Choice’ Test on Mean Values of Travel Time Savings

(i) Choice set level						
<i>\$/person hour (VTTS):</i>	Running cost					
	All attributes relevant			Allowing for attribute being ignored		
	Full	Plausible	Difference	Full APS	Plausible APS	Difference
Free flow time	\$13.01	\$12.53	3.81%	\$12.02	\$11.62	3.44%
Slowed down time	\$13.93	\$13.85	0.58%	\$14.52	\$14.53	-0.07%
Trip time variability	\$2.57	\$2.53	1.58%	\$2.33	\$2.95	-21.02%
<i>\$/person hour (VTTS):</i>	Toll cost					
	All attributes relevant			Attribute processing strategy applied		
	Full	Plausible	Difference	Full APS	Plausible APS	Difference
Free flow time	\$10.16	\$10.51	-3.33%	\$9.08	\$9.73	-6.68%
Slowed down time	\$10.88	\$11.61	-6.29%	\$10.96	\$12.17	-9.94%
Trip time variability	\$2.00	\$2.12	-5.66%	\$1.76	\$2.47	-28.75%
Weighted average VTTS:	12.49	12.29	1.63%	11.84	11.83	0.09%

(ii) Respondent level						
<i>\$/person hour (VTTS):</i>	Running cost					
	All attributes relevant			Allowing for attribute being ignored		
	Full	Plausible	Difference	Full APS	Plausible APS	Difference
Free flow time	\$13.01	\$13.08	-0.54%	\$12.02	\$13.37	-10.10%
Slowed down time	\$13.93	\$15.06	-7.50%	\$14.52	\$16.89	-14.03%
Trip time variability	\$2.57	\$3.47	-25.94%	\$2.33	\$3.01	-22.59%
<i>\$/person hour (VTTS):</i>	Toll cost					
	All attributes relevant			Attribute processing strategy applied		
	Full	Plausible	Difference	Full APS	Plausible APS	Difference
Free flow time	\$10.16	\$11.57	-12.19%	\$9.08	\$10.73	-15.38%
Slowed down time	\$10.88	\$13.33	-18.38%	\$10.96	\$13.56	-19.17%
Trip time variability	\$2.00	\$3.07	-34.85%	\$1.76	\$2.42	-27.27%
Weighted average VTTS:	10.34	12.01	-13.91%	9.55	11.44	-16.52%

3.2 Pairwise Alternative ‘Plausible Choice’ Test and the Presence of Dominance

So far we have not discussed the possibility of dominance and what role it might play as an embedded feature of the design, as well as a response issue. We need to introduce some definitions related to dominance in order to be clear as to what features of the choice experiment setting we are investigating. Dominance has two possible interpretations. The first, which is more common, relates to *design issues* where one would distinguish between (i) the choice set level, where an alternative is equal or better on all attributes to another alternative, and (ii) across all choice sets shown to a respondent, where a particular alternative is always better on all attributes (noting this never happens in our choice experiment designs). The second relates to *response issues* where (i) at the choice set level, the respondent chooses an alternative which is best on all attributes, and (ii) at the respondent level across all choice sets, where a particular alternative is always chosen regardless of whether it is always the best in each choice set. We focus on the second interpretation, but report the extent of dominance (as a design issue) in the design being used. It must be noted, however, that design and response issues are not independent. In particular, the presence of design dominance allows a respondent in a single choice scenario to choose an alternative that is equal or inferior on all attributes to another alternative. One reason for this might be a preference for or against an alternative, where this might lead to all choice responses being made for that alternative or class of alternatives (e.g., status quo (SQ) or not status quo). If we considered SQ to be an attribute of the alternatives, then the preference for or against SQ might break the dominance condition. However we do not know a priori what the sign will be for any one respondent.

A weaker plausibility test compares the pairs of alternatives, allowing a choice to be considered consistent with a number of plausible heuristics such as elimination-by-aspects (EBA) even when it contains no best attributes, if it has at least one better attribute than the rejected alternative on a pairwise comparison. If the pair includes the reference alternative, it may be that this contrast delivers an outcome that passes a pairwise ‘plausibility choice’ test on more occasions.

On closer inspection, of the 54 choice sets that failed the full choice set ‘plausible choice’ test from the 6,048 choice sets in the sample, all but one satisfied the pairwise ‘plausible choice’ test, with 20 of the chosen alternatives having the better level on all five 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 ‘plausibility choice’ 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?⁷ For modelling, it may be appropriate to remove the reference alternative and treat their

⁷ Within the environmental economics literature this is actually an often quoted criticism of eliciting preferences through stated preference methods (i.e., that people act strategically in an hypothetical setting and are more likely to choose a non reference as it provides them with an “option” to choose it, even though they would be unlikely to do so in reality). Related to this issue of strategic decision making is yeah-saying (especially in environmental economics case-studies). Within the context of the transport application herein, this is far less likely to be of concern; however it is important to recognise this matter in applications more aligned to environmental economics.

processing strategy as elimination by alternatives, allowing the reference alternative to be specified as ‘non-existent’. This is equivalent to ignoring 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.

At a choice set level, if a chosen 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 to or worse than the other alternative. While the pairwise ‘plausible choice’ test applied above to those who failed the three-way ‘plausible choice’ test found only one case of response dominance at the choice set level, an examination of *all choice sets for each respondent* uncovered a wider pattern of choice of a dominated alternative for 46 (out of 6,048) observations. Of the total of 6,048 choice scenarios, 667 contained a dominated alternative⁸, meaning that 6.9 percent of choice tasks containing dominance led to the choice of a dominated alternative. The 46 cases of dominated alternatives being chosen are summarised in Table 6 (where the focus is on response and design dominance). 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 (e.g., there were 10 choice observations where the reference alternative was best on all attributes but not chosen, in contrast there were 28 choice observations where SC alternative 3 was best but not chosen). Of note is this high number for alternative three, stemming from choice scenario 20 in the experimental design (see Appendix A), where the reference alternative was inferior, on the presented attributes, to the third alternative. 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. In particular, the third alternative might have been missed because most of the two hypothetical alternatives had a toll attached, whereas the reference alternative never did⁹. Those respondents who placed greater disutility on a toll might have disengaged from the hypothetical alternatives, or considered all of them as toll road alternatives. If this is the case, then it is likely that this phenomenon is occurring in other choice scenarios, when dominance is not present, to the detriment of the quality of the dataset. Care should be taken to minimise the chance of this happening, via clear instructions to the respondent and, if relevant, appropriate training of the interviewers administering the survey.

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

⁸ The 667 choice scenarios containing dominance primarily stemmed from three choice scenarios containing dominance in the experimental design (see Appendix A, choice scenarios 15, 20 and 25). However, the application of various rules to ensure variation in the attribute levels of the hypothetical alternatives might have led to the presentation and capture of choice scenarios containing dominance that was not present in the experimental design.

⁹ The experimental design did not contain a scenario where the second alternative dominated the reference alternative. However, the application of various rules as outlined in the previous footnote led to this condition in some of the choice scenarios in the dataset.

experiment, while the two hypothetical 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, shown in the last column of Table 6, the respondent always chose the reference alternative over the 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. 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 implausible 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.

The above examination of dominance assumed that none of the attributes were ignored. Just as the number of alternatives in a dataset that lead to failure in the 'plausible choice' test will be impacted by the particular attribute processing strategy of the individual, so too will the presence of dominance in a choice task. If an alternative is already dominated by another alternative, then the omission of attributes in the comparison will either retain the dominance or lead to a tie between the two alternatives. However, a pair of alternatives that, under full attribute attendance, present tradeoffs, with some attributes better and worse for each alternative, might degenerate into a condition where one alternative dominates the other. Choice of a dominated alternative in this scenario might be indicative of several things. A genuine mistake might have been made either at the time of choice or when revealing which attributes were ignored. Alternatively, the attribute processing rules might vary across choice tasks, even though they were gathered once after the completion of the choice scenarios in this study (see Puckett et al. 2007 for a study where attribute processing strategies were collected after every choice scenario). A consequence of this condition is that even when a design is generated that has no dominance when full attribute attendance is assumed, the choice scenarios might appear to the respondent to contain dominance when their specific attribute processing strategy is taken into account, and this might have implications for the effectiveness and/or efficiency of the design. A potentially very important area of future research is the design of stated choice experiments that are robust to a mix of attribute processing strategies.

Table 6: Response Dominance in the Full Sample

Number of observations for which each alternative dominated the chosen alternative		Choice behaviour over all 16 tasks for those choice observations that were dominated	
Reference	10	Always chose reference alternative	9
SC Alternative 2	7	Never chose reference alternative	10
SC Alternative 3	28	Mix of reference and other alternatives	27
Reference and SC Alternative 2	1	-	-
Total	46	Total	46

3.3 Influences of 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) at the respondent level, where the binary dependent variable equals 1 for 23 observations who always choose the reference alternative across all 16 choice sets, 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 had expected to 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 being ignored by the respondent (the latter obtained from supplementary questions). This latter evidence might be due the presence of greater engagement in evaluating the new alternatives. Also, with greater

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 2 with extra influences	Base model 2 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

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. A closer look at the data confirms 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 expanded on the binary choice base Model 2 to accommodate the full set of three alternatives under full attribute relevance (Model 3) and under attribute non-preservation (Model 4). 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. The Bayes Information Criterion (BIC), which accounts for the number of parameter estimates, judges a model by how close its fitted values tend to be to the true values, in terms of a certain expected value (Akaike 1974). The BIC value assigned to a model *ranks* competing models and indicates which is the best among the given alternatives, and is preferred over the log-likelihood criteria when the number of parameters changes (as a way of avoiding overfitting). Model 4 with a lower BIC, is a significant improvement over Model 3.

3.4 Dimensional versus Holistic Processing Strategies

Another pairwise test could be based on the ‘majority of confirming dimensions’ (MCD) rule (Russo and Doshier 1983), which is concerned with the total count of superior attributes in each alternative. Under this test, pairs of attributes are compared in turn, with an alternative winning if it has a greater number of better attribute levels. 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 allowed¹⁰. The distribution of the number of best attributes is shown in Table 8, both for the full relevance sample, and accounting for attributes being ignored, with separate reporting for all alternatives and the chosen alternative only. The distribution for the chosen alternative is skewed towards a higher number of best attributes in both cases, and higher means can also be observed, which is plausible. 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 relative utilities.

A close inspection shows that the percentage of alternatives with zero strictly best attributes is much higher when allowing for attributes being ignored than in the ‘full relevance’ group (compared to the other rows of evidence). This might suggest that respondents are more likely to ignore an attribute when at least one attribute is outranked. On this evidence, if found true in other data sets, it has important behavioural implications since the analyst may wish to remove alternatives in model estimation where the number of best attributes is zero.

¹⁰ Accounting for ties did not materially affect the findings.

Table 8: Number of Strictly Best Attributes per Alternative

Number of best attributes	Full relevance				Allowing for attribute being ignored			
	All alternatives		Chosen alternative		All alternatives		Chosen alternative	
	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	

The model results are reported in Table 9, with Model 1 representing the base model, with all attributes assumed to be considered. Model 2 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 Bayes Information Criterion (BIC) which accounts for the number of parameter estimates can be observed.

Model 3 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.

The same tests were performed, after accounting for attributes stated as being ignored (Models 4 to 6). Any ignored attributes were not included in the count of the number of best attributes. Model 4 of Table 9 sets out the base model that accounts for attribute ignoring, which itself fits the data better than when all attributes are assumed to be attended to. Model 5 presents the model that accounts for both heuristics, and Model 6 represents the inclusion of the number of best attributes in the absence of explicit consideration of each attribute, after allowing for attributes that are indicated as ignored. The BIC is improved, at 1.7483 compared to 1.7514 for the base model, with the number of best attributes parameter being statistically significant and of the expected sign.

We report the weighted average VTTS in Table 9 (where the weights are the levels of each attribute, namely free flow and slowed down time, and running and toll cost) which, at the mean estimate for the weighted average total time, appear to vary sufficiently between full relevance and allowing for attributes being ignored, but not between models within each of this attribute processing settings when allowance is made for the number of attributes that are best. When confidence intervals are generated using a bootstrapping procedure with 1,000 random draws from normal distributions for relevant parameters, with moments set at their coefficient point estimates and standard errors (Krinsky and Robb 1986), we find, as expected that there are no statistically significant differences between Models 1 and 2 (and between Models 4 and 5); however the differences are statistically significant at 95 percent confidence level between the estimates for full relevance and attribute non-attendance.

Table 9: Influence of Majority of Confirming Dimensions

	Full Relevance			Allowing for attributes being ignored		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Base	Base plus # best attributes	# best attributes	Base	Base plus # best attributes	# best attributes
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)
<i>Value of travel time savings (\$/person hour):</i>						
Free flow time (based on running cost parameter estimate)	13.01	13.22		12.03	12.18	
Free flow time (based on toll cost parameter estimate)	10.15	9.70		9.08	8.66	
Slowed down time (based on running cost parameter estimate)	13.93	12.80		14.52	13.55	
Slowed down time (based on toll cost parameter estimate)	10.88	9.40		10.96	9.63	
Weighted average VTTS:	12.48	12.20		11.85	11.58	
<i>Number of observations with attribute ignored:</i>						
Free flow time	-			944		
Slowed down time	-			1504		
Trip time variability	-			2240		
Running cost	-			1120		
Toll cost	-			656		
<i>Model fit:</i>						
Log-likelihood at convergence	-5428.17	-5417.55	-6224.89	-5265.81	-5252.05	-6123.98
BIC	1.8051	1.8031	2.0628	1.7514	1.7483	2.0295
Sample size	6048					

While Model 2 (Model 5) compared to Model 1 (Model 4) is an improvement on BIC, albeit relatively small, its underlying form suggests that all respondents simultaneously consider and trade between both the attribute levels in a typical compensatory fashion (both under full relevance and after ignoring some attributes if applicable), 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, two latent class models¹¹ were estimated (Table 10). Two classes are defined¹², where the utility expressions in each class are constrained

¹¹ See Hensher and Greene (2010) for other examples of the identification of attribute processing heuristics with the latent class model.

¹² 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

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 BIC under full attribute relevance (and accounting for ignored attributes) improving from 1.8051 (1.7514) for the base model, to 1.8031 (1.7483) for the single class model that contains both the levels and the number of best attributes, to 1.7795 (1.7287) for the latent class model. Again the number of best attributes parameter is statistically significant and of the expected sign.

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 80 percent for processing of the constituent attributes and between 15 and 18 percent for the number of attributes being the determining influence.

Table 10: Identifying Role of MCD - Latent Class Model

	Full Relevance	Allowing for attributes being ignored
<i>Class 1</i>		
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)
<i>Class 2</i>		
# of attributes in an alternative that are best	0.2856 (2.76)	0.2665 (3.06)
<i>Probability of class membership:</i>		
Class 1	0.8465 (6.25)	0.8206 (9.58)
Class 2	0.1535 (6.35)	0.1794 (8.17)
<i>Value of travel time savings (\$/person hour):</i>		
Free flow time (based on running cost parameter estimate)	13.54	12.45
Free flow time (based on toll cost parameter estimate)	10.34	9.61
Slowed down time (based on running cost parameter estimate)	13.08	14.41
Slowed down time (based on toll cost parameter estimate)	9.99	11.13
<i>Weighted average VTTS:</i>	<i>12.60</i>	<i>12.17</i>
<i>Number of observations with attribute ignored:</i>		
Free flow time (mins)	-	944
Slowed down time (mins)	-	1504
Trip time variability (plus/minus mins)	-	2240
Running cost (\$)	-	1120
Toll cost (\$)	-	656
<i>Model fit:</i>		
BIC	1.7795	1.7287
Log-likelihood at convergence	-5402.47	-5218.52
Sample size	6048	

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.

The implication is that the application of the choice model must recognise that the trading amongst the attributes occurs up to a probability of 85 percent (or 82 percent when accounting for ignoring) on average, with the number of best attribute levels having an influence up to a probability of 15 percent (or 18 percent) 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. When we compare the mean estimates of VTTS for Model 2 (and Model 5) in Table 9 with the latent class models, the mean estimates are respectively \$12.20 and \$12.60 for full relevance and \$11.58 and \$12.17 when attributes are ignored. The latent class mean estimates have moved closer to the mean estimates in Table 9 when we do not include allowance for the number of best attributes (i.e., Model 1 and 4 in Table 9 of \$12.48 and \$11.85 respectively). If the contrast is with the base models in Table 9, we would conclude that the VTTS estimates are not statistically significant in the presence and absence of accounting for the ‘majority of confirming dimensions’ rule; however differences are significant when allowing for attributes to be ignored. This finding supports the evidence in studies undertaken by Hensher and his colleagues (see Hensher 2010) that allowing for attribute non-attendance has a statistically significant influence on the mean estimates of VTTS.

3.5 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 a reference alternative attribute level was better, equal and worse relative to SC1 and SC2, defined as a series of attribute specific dummy variables (e.g., free flow time (FFT) better = 1 if reference FFT minus SC1 FFT is negative and equal to zero if reference FFT minus SC1 FFT is positive). The choice response variable refers to the alternative chosen. A simple logit model was specified in which we included the better and worse attribute forms for all five design attributes (eliminating ‘worse’ for toll cost only because 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 as illustrated above for FFT), 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 (although marginally significant). 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

Attributes defined as Reference minus SC1 or minus 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	

3.6 Revision of the Reference Alternative as Value Learning

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). There is an important distinction to be made between value learning, which in its broadest meaning implies underlying preferences are changing, and reference revision which 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*. The latter is a special case of the former. We focus on value learning.

We ran a model in which we identified the chosen alternative from a previous choice set, and created a dummy variable equal to 1 associated with whatever alternative was chosen in the previous choice set, be it the initial reference alternative or one of the offered non-status quo alternatives (namely alternatives two or three). We then introduced into the utility expressions the revised reference dummy variable 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) for this variable, which suggests that when the reference alternative is revised, in the next choice scenario it 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 only through a correlated error variance specification, where the latter captures many unobserved effects at the alternative level.

Table 12: Identifying Role of Value Learning

Note: Choice set 1 is removed

	Full Relevance
<i>Revised Reference (1,0)(which can be any of the three alternatives)</i>	<i>0.9358 (15.73)</i>
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

3.7 A Revised Model for Future Stated Choice Model Estimation

We present a model below as a contrast to the base model (Table 9, Models 2 and 5), where we include value learning, the majority of confirming dimensions, and attribute non-attendance. This model captures a main contribution of this paper. 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 set 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.

The weighted mean estimate of value of travel time savings in Table 13 is \$11.19 per person hour. This estimate 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 \$12.49 under full attribute reference, or \$11.84 when we allowed for attributes being ignored. At the 95 percent level of confidence, the weighted mean estimate of VTTS is significantly different and lower.

Table 13: Revised Full Model for Future 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)
Value learning reference revision (1,0) which may be the original reference alternative	0.8843 (13.8)
Initial Choice Set Reference dummy (1,0) for choice set 1	1.1442 (8.99)
BIC	1.6092
Log-likelihood at convergence	-4600.45
Sample Size	5793
<i>Mean Value of Travel Time Savings (\$/person hour):</i>	
Free flow time (based on running cost parameter estimate)	12.02
Free flow time (based on toll cost parameter estimate)	8.50
Slowed down time (based on running cost parameter estimate)	12.40
Slowed down time (based on toll cost parameter estimate)	8.77
<i>Weighted average VTTS:</i>	11.19

4 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, 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 'plausible choice' test proposed herein for the entire choice set, and for pairwise alternatives, at the observation and respondent levels, 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 implausible choice behaviour stems from the recognition that the choice might be plausible 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 the respondent claimed not to have considered, the application of the MCD heuristic, and revision of the reference alternative as value learning. 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 implausible might pass the plausibility 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, plausibility 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 implausible when viewed through the conventional prism of utility maximisation.

Of interest to the analyst are possible ways in which implausible 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 implausible behaviour, which suggests that the number of choice tasks might not have an impact, within reasonable limits. Choice task complexity (as defined by dimensions such as number of alternatives, attributes and attribute levels) was not varied in this analysis; however the impact of task complexity on implausible behaviour would be an interesting area of research. Also of interest is the plausibility 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 plausibility 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 plausible choice in the stated choice environment might not actually be the best way forward. Survey realism might instead be more important.

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 appear to provide plausible explanations of utility maximising behaviour in choice making. Testing of the ideas presented on other data sets will enable us to establish the portability of the evidence.

Acknowledgements

We thank John Rose, Joffre Swait, Sean Puckett, Bill Greene, Victor Adamowitz, Stephane Hess and four referees for very useful detailed comments on earlier versions.

Appendices

Appendix A

Alt.	Cset	Block 1					Cset	Block 2				
		FF	SDT	Var	Cost	Toll		FF	SDT	Var	Cost	Toll
1	1	0	0	0	0	0	17	0	0	0	0	0
2		-0.15	0.3	-0.3	0.3	3.5	0.15	0.3	-0.3	-0.3	0.5	
3		0.15	-0.3	-0.3	-0.3	0	0.15	-0.15	-0.15	0.3	3	
1	2	0	0	0	0	0	18	0	0	0	0	0
2		-0.15	-0.3	0.3	-0.15	4	0.15	-0.15	-0.15	-0.15	2.5	
3		0.3	-0.15	-0.3	0.15	1	0.15	-0.15	0.15	-0.3	1.5	
1	3	0	0	0	0	0	19	0	0	0	0	0
2		0.3	-0.15	-0.15	-0.15	0	-0.3	0.15	0.15	0.3	0	
3		-0.15	-0.3	-0.3	-0.3	0.5	0.3	-0.15	-0.3	0.3	0.5	
1	4	0	0	0	0	0	20	0	0	0	0	0
2		-0.3	-0.15	0.3	-0.3	3.5	0.3	0.3	-0.3	-0.15	0.5	
3		0.3	-0.3	0.3	-0.3	2.5	-0.15	-0.3	-0.15	-0.15	0	
1	5	0	0	0	0	0	21	0	0	0	0	0
2		-0.15	-0.15	-0.3	-0.3	2.5	-0.15	0.15	0.3	-0.3	3	
3		-0.3	-0.15	-0.3	-0.15	3	0.3	0.15	0.3	-0.3	2.5	
1	6	0	0	0	0	0	22	0	0	0	0	0
2		0.3	0.15	-0.3	-0.3	3.5	-0.15	0.3	-0.3	0.3	1.5	
3		-0.15	-0.3	-0.15	0.15	3.5	-0.15	-0.3	-0.15	-0.15	2	
1	7	0	0	0	0	0	23	0	0	0	0	0
2		-0.3	0.15	0.15	-0.15	1.5	0.3	0.15	-0.15	-0.3	0	
3		-0.3	0.3	-0.15	0.15	3	-0.3	0.3	0.3	-0.15	4	
1	8	0	0	0	0	0	24	0	0	0	0	0
2		-0.15	-0.3	-0.3	-0.15	2.5	0.3	-0.3	0.3	0.15	0.5	
3		-0.3	-0.15	0.3	0.15	2	0.15	0.3	0.15	-0.15	2.5	
1	9	0	0	0	0	0	25	0	0	0	0	0
2		-0.3	-0.3	-0.15	0.15	2	-0.3	-0.3	-0.3	0.3	0	
3		0.15	-0.15	-0.3	-0.3	4	-0.3	-0.15	0.3	0.3	4	
1	10	0	0	0	0	0	26	0	0	0	0	0
2		-0.15	0.3	-0.15	0.15	4	-0.3	-0.15	0.15	0.3	1	
3		0.3	-0.15	0.15	-0.3	1	-0.3	0.15	0.3	-0.3	1.5	
1	11	0	0	0	0	0	27	0	0	0	0	0
2		0.15	-0.3	0.15	-0.3	4	0.15	-0.15	-0.15	-0.15	3	
3		-0.15	0.3	-0.3	0.15	2	-0.15	0.15	-0.15	-0.15	3.5	
1	12	0	0	0	0	0	28	0	0	0	0	0
2		-0.15	0.15	0.3	-0.3	2	-0.3	-0.3	-0.15	0.15	4	
3		-0.3	-0.3	0.15	-0.3	3.5	-0.15	0.15	0.15	-0.15	1.5	
1	13	0	0	0	0	0	29	0	0	0	0	0
2		-0.15	-0.15	0.3	-0.15	1.5	0.15	-0.15	-0.15	0.15	0	
3		-0.3	-0.15	-0.3	-0.15	4	-0.15	0.15	-0.3	0.3	0.5	
1	14	0	0	0	0	0	30	0	0	0	0	0
2		0.15	-0.3	-0.15	-0.15	1	0.3	-0.15	-0.15	0.15	1	
3		0.3	0.15	-0.15	-0.3	1	-0.3	-0.3	-0.15	0.3	3.5	
1	15	0	0	0	0	0	31	0	0	0	0	0
2		-0.15	-0.3	0.15	-0.15	0.5	-0.3	0.3	-0.3	0.3	3	
3		0.15	-0.3	0.15	0.15	3	-0.15	0.3	-0.15	0.3	0.5	
1	16	0	0	0	0	0	32	0	0	0	0	0
2		-0.3	-0.3	-0.3	-0.3	2	-0.3	-0.15	0.15	-0.3	3	
3		-0.15	0.3	-0.15	-0.15	0	-0.3	-0.3	-0.3	-0.15	3.5	

Columns include the alternative number (Alt.), where alternative 1 is the reference alternative, a unique choice scenario identifier (Cset), freeflow time (FF), slowed down time (SDT), trip time variability (Var), running cost (Cost) and toll (Toll). FF, SDT, Var and Cost are all expressed as a proportion of the recent trip values, where this proportion is added to the recent trip value. Toll is in dollars. Grey highlighting denotes the best attribute level in a choice scenario. Alternatives marked in bold denote a situation of dominance under full attribute attendance.

Appendix B

Alt.	FF	SDT	Var	Cost	Toll	TotT	TotC	Choice
1	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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

Alt.	FF	SDT	Var	Cost	Toll	TotT	TotC	Choice
1	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	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	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	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	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	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	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	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	20	20	10	4.16	0	40	4.16	0
2	14	26	7	5.41	3	40	8.41	0
3	17	26	8	5.41	0.5	43	5.91	1
1	80	10	15	11.18	0	90	11.18	0
2	56	13	10	14.53	3	69	17.53	0
3	68	13	13	14.53	0.5	81	15.03	1
1	60	10	12	8.58	0	70	8.58	0
2	42	13	9	11.15	3	55	14.15	0
3	51	13	11	11.15	0.5	64	11.65	1
1	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	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	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	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	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	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	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	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	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

Alt.	FF	SDT	Var	Cost	Toll	TotT	TotC	Choice
1	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	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	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	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

Grey highlighting denotes the best attribute level in a choice scenario.

References

- Akaike, H. 1974. A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19 (6) 716-723.
- Bateman, I. J., R. T. Carson, B. Day, D. Dupont, J. J. Louviere, S. Morimoto, R. Scarpa and P. Wang (2008) Choice set awareness and ordering effects in discrete choice experiments, CSERGE Working Paper EDM 08-01.
- Cameron, T. A. and J. R. DeShazo (2010) Differential attention to attributes in utility-theoretic choice models, *Journal of Choice Modelling*, 3 (3) 73-115.
- 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 J. L. Pinto (2010) Ordering anomalies in choice experiments, *Journal of Environmental Economics and Management*, 59 (3) 271-285.
- Day B., I. J. Bateman, R. T. Carson, D. Dupont, J. J. Louviere, S. Morimoto, R. Scarpa and P. Wang. 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 (3) 360-385.
- Fiebig, D., M. Keane, J. J. Louviere and N. Wasi. 2010. The generalized multinomial logit: accounting for scale and coefficient heterogeneity, *Marketing Science*, 29 (3) 393-412.
- Gilbride, T. J. and G. M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules, *Marketing Science*, 23 (3) 391-406.
- Greene, W. H. and D. A. Hensher. 2010. Does scale heterogeneity matter? A comparative assessment of logit models, *Transportation*, 37 (3) 413-428.
- Haaiker, R., K. Wagner and M. Wedel. 2000. Response latencies in the analysis of conjoint choice experiments, *Journal of Marketing Research*, 37 (3) 376-382.
- Hensher, D. A. 2010 Attribute processing, heuristics and preference construction in choice analysis, in Hess, S. and A. Daly (eds.) *State-of Art and State-of Practice in Choice Modelling*, Emerald Press, U.K., 35-70.
- Hensher, D. A. and W. H. Greene. 2010. Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification, *Empirical Economics*, 39 (2) 413-426.

- Hensher, D. A., W. H. Greene and C. Chorus. 2010. Random regret minimisation or random utility maximisation: an exploratory analysis in the context of automobile fuel choice, submitted to *Journal of Advanced Transportation*.
- Hess, S. and D. A. Hensher. 2010. Using conditioning on observed choices to retrieve individual-specific attribute processing strategies, *Transportation Research Part B*, 44 (6) 781-90.
- Kahneman D. and A. Tversky. 1979. Prospect theory: an analysis of decisions under risk, *Econometrica*, 47 (2) 263-291.
- Krinsky I. and A. L. Robb. 1986. On approximating the statistical properties of elasticities, *The Review of Economics and Statistics*, 68 (4) 715-719.
- Lancsar, E. and J. J. Louviere. 2006. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences?, *Health Economics*, 15 (8) 797-811.
- Louviere, J. J., D. A. Hensher and J. D. Swait. 2000. *Stated Choice Methods and Analysis*, Cambridge University Press, Cambridge.
- McNair, B. J., J. Bennett and D. A. Hensher. 2010a. A comparison of responses to single and repeated discrete choice questions, *Resource and Energy Economics*, published on line <http://dx.doi.org/10.1016/j.reseneeco.2010.12.003>
- McNair B. J., D. A. Hensher and J. Bennett. 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.
- Puckett, S., D. A. Hensher, J. M. Rose and A. Collins. 2007. Design and development of a stated choice experiment in a two-agent setting: interactions between buyers and sellers of urban freight services, *Transportation*, 34 (4) 429-451.
- Rose, J. M. and M. C. J. Bliemer. 2008. Stated preference experimental design strategies, in Hensher, D. A. and K. J. Button (eds.) *Handbook of Transport Modelling*, Oxford, Elsevier, 151-179.
- Rose, J. M. and I. Black. 2006. Means matter, but variance matter too: decomposing response latency influences on variance heterogeneity in stated preference experiments, *Marketing Letters*, 17 (4) 295-310.
- Russo, J. E. and B. A. Doshier. 1983. Strategies for multiattribute binary choice, *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 9 (4) 676-696.
- Scarpa, R., D. Campbell and W. G. Hutchinson. 2007. Benefit estimates for landscape improvements: sequential bayesian design and respondents' rationality in a choice experiment study, *Land Economics*, 83 (4) 617-634.
- Swait, J. D. 2009. Choice models based on mixed discrete/continuous PDFs, *Transportation Research Part B*, 43 (7) 766-783.