Identifying a behaviourally relevant choice set from stated choice data

David A. Hensher · Chinh Ho

Abstract Stated choice experiments are designed optimally in a statistical sense but not necessarily in a behavioural choice making sense. Statistical designs, and consequently model estimation, assume that the set of alternatives offered in the experiment are processed by respondents with a specific processing strategy. Much has been studied about attribute processing using discrete choice methods in travel choice studies, but this paper focuses more broadly on processing of alternatives in the choice set offered in the experiment. This paper is motivated by the primary idea that the distribution of predicted choice probabilities associated with a set of alternatives defining a given choice set might provide strong evidence on the strategies that agents appear to use when choosing a preferred alternative. In an empirical setting of a choice set of size three, four model specifications are considered including a model for the selection of the best alternative in the full choice set and three variants of a best–worst regime. Using state choice data on road pricing reform, the empirical analysis examines which model specification delivers the most accurate prediction of the chosen alternative. The results suggest which alternatives really matter in choice making and hence the alternatives that might be included in a choice set for model specification.

Keywords Best–worst · Model specifications · Choice models · Choice sets · Full rank · Irrelevance of Statewise Dominated Alternatives · Road pricing reform · Choice experiment

D. A. Hensher (✉) · C. Ho
Institute of Transport and Logistics Studies (ITLS), The University of Sydney Business School, The University of Sydney, Camperdown, NSW 2006, Australia
e-mail: David.Hensher@sydney.edu.au

C. Ho
e-mail: chinh.ho@sydney.edu.au

🖄 Springer

Introduction

Stated choice experiments have proven popular as a framework to seek respondent preferences for various alternatives, especially those that are not currently available in the market. With discrete choice methods, respondent preferences are 'revealed' through their responses to a number of scenarios, also known as choice tasks. Each choice task may ask respondents to choose the most preferred alternative, provide a full rank ordering of all alternatives offered, or choose the most and least preferred options (best–worst see Marley and Flynn, in press). Within the setting of utility maximisation (or regret minimisation), these various response metrics are all candidate derivatives of the full rank order.

It has become common practice in stated choice studies to preserve all offered alternatives in a choice set when it comes to model estimation. This includes the recent development of exploded best–worst methods which use a rank order of alternatives in each choice task to create pseudo observations that are then added to the dataset, hence the term exploded. One way to create pseudo observations from full rank order (best–worst) data is to remove the most preferred (best) alternative from each choice task, redefine the least preferred (worst) alternative as the chosen alternative for that choice task and reverse the attribute sign (i.e., replace X with $-X$) (Rose 2014). The selection of the number of alternatives in a choice set for model estimation is typically based on some arbitrary assumption linked to notions of comprehendability of the choice experiment.

Although we cannot formally identify the optimal number of alternatives that the respondents might consider and the way in which they process the alternatives presented,[1] we can attempt to gain a better understanding of the extent to which all alternatives are value adding in preference revelation. Conversely, we can investigate whether a subset of the alternatives is all that really matters for an individual when we develop a practical model to predict the chosen (most preferred) alternative. This raises many questions of behavioural curiosity, including whether all alternatives shown in a stated choice experiment should be preserved for each sampled respondent in estimation, or whether some other paradigm of choice set formation might be more reflective of a processing rule adopted by the respondents. This paper is aligned to the broader literature on context effects which examines how preference changes depend on the availability of other options. The context effects include compromise, attraction and similarity (see Rieskamp et al. 2006 and Trueblood et al. 2013; Rooderkerk et al. 2011, Leong and Hensher 2012a, b, 2014 for details). Our research begins with recognising the nature of stated choice (SC) experiments that typically predefine a set of alternatives offered in a choice scenario. As such, choices are partly driven by the context provided by the set of 'designed' alternatives. Given the constructed rank order data collected through a best–worst survey, we investigate the role that a specific alternative might have in influencing respondent preferences using a number of testable model specifications (as implied processing strategies, detailed in the following sections). We are interested in establishing which model specification supports the best performance of the model in terms of predicting the most preferred alternative as chosen in each choice task.

The paper is structured as follows. We begin by presenting a number of model specifications designed to represent how an individual might consider or ignore an alternative such that the most preferred alternative in each of the four choice tasks has the highest

---

[1] Specifically, in model estimation and application, in contrast to the design of a choice experiment.

predicted choice probability (referred to as the 'compliance test'). We draw on a choice experiment to study car user preferences for alternative road pricing reform packages as a way of empirically identifying the incidence of compliance. The findings obtained from a series of mixed logit models are then presented, followed by interpretation and an inquiry into socioeconomic and contextual factors that may explain, in part, the relative incidence of compliance across the model specifications. The paper concludes with a summary of the main findings and what this means for ongoing discrete choice applications. We include an Appendix (2) as a theoretical perspective offered in the broader literature on choice and uncertainty. The underlying theory is centred on the Axiom of Irrelevance of Statewise Dominated Alternatives (ISDA) which provides an appealing basis for testing the compliance with utility maximisation of each of the model specifications.[2]

Alternative model specifications

Our focus on choice set processing begins with two sequenced choice response questions that elicit the best and worst alternatives in a three-alternative choice scenario. This is a typical forced choice response with no invitation to reject any alternatives through a null choice response. With only three alternatives, the analyst can construct full rank data from such responses. There are a number of possible choice sets that an individual might process depending on the relevance of all or a subset of alternatives presented in a choice scenario. A choice set containing all alternatives would typically assume that respondents consider all offered alternatives when indicating their preferences. Apart from this full choice set, other possible choice sets can be formed for model estimation, depending on an a priori processing strategy adopted by each individual. Given the full rank order of three alternatives considered in this paper,[3] there are at least four candidate modelling specifications (MSs) that describe processing strategies worthy of consideration:

1. MS1: consider the best alternative in the full choice set of 3 alternatives;
2. MS2: an exploded best–worst choice set. This specification considers the best alternative in the full choice set and the worst alternative of the remaining two options—see further explanation below;

---

[2] A referee suggested that the ISDA section in the original paper would be better presented as an Appendix in order to not detract from the main contribution of the paper. Given our interest in identifying the extent to which respondents, under each model specification, preserve their preference ordering of alternatives, it can be studied without any reference to ISDA; however we argue that the inclusion of the ISDA context offers an appealing setting in which to judge compliance of preferences with a behaviourally interesting axiom. Hence it is set out in Appendix 2.

[3] While the actual survey instrument (i.e., choice scenarios) asked for best and worst alternatives (both questions shown on the screen at the same time), the fact that we have only 3 alternatives allows us to construct a full preference ranking dataset. What we are doing in this paper, just like the inferential literature on attribute non-attendance (see Hess and Hensher 2010), is to use the information from two choice questions to construct a full rank response to see if we can establish (also guided by ISDA—see Appendix 2) what rule might be being used (in terms of how the alternatives are processed) to choose an alternative from an offered choice set of up to 3 alternatives, up to a predicted probability. There is nothing in the full rank data constructed by the analyst to suggest that all three alternatives were actually taken into account as acceptable alternatives in arriving at the best or worst alternative; indeed the choice may have been associated with only two alternatives. This seems to us to be a reasonable and very interesting way of understanding what alternatives in a choice set imposed by a stated choice experiment are actually relevant (as per arguments developed elsewhere by Hensher 2006).

3. MS3: consider the best of the most and second most preferred alternatives; and
4. MS4: consider the best of the most and least preferred alternative.

The four models associated with the four processing rules differ mainly in the number of alternatives and which alternatives are being used to form a choice set. However, they are not different in terms of utility specification.

We can align these MSs to the best–worst literature and use Marley and Flynn's (in press) notation, to express the four models as follows.[4] Let Y denote a full choice set with 3 alternatives, $\rho = \rho_1, \rho_2, \rho_3$ be a typical rank order of the alternatives in Y from most preferred (best; ρ_1) to least preferred (worst; ρ_3). Note that the respondent is only asked to indicate the best and worst alternatives (in two questions) and is not forced to rank the three alternatives (the full ranking being constructed through simple deduction by the authors). Given the two questions asked to respondents, MS3 corresponds to a behaviourally appealing processing rule where respondents implicitly chose between the best and the second best and ignore the worst. Similarly, MS4 assumes that respondents processed the best and the worst alternatives (for) while ignoring the second best alternative. In selecting a particular alternative as the worst they might be actually rejecting that alternative as relevant, even though they are 'forced' to indicate the worst alternative.[5] The four modelling assumptions are expressed mathematically as follows (noting ρ_1 is best and ρ_3 is worst, $B_Y(y)$ denotes the probability that alternative y is chosen as best in Y, and $W_Y(y)$ is the probability that alternative y is chosen as worst in Y)[6]:
MS1:

$$B_Y(\rho_1) = \frac{\exp(\beta X_1)}{\sum_{j=1,2,3} \exp(\beta X_j)}$$

MS2:

$$B_Y(\rho_1) W_{Y-\{\rho_1\}}(\rho_3) = \frac{\exp(\beta X_1)}{\sum_{j=1,2,3} \exp(\beta X_j)} \frac{\exp(-\beta X_3)}{\sum_{j=2,3} \exp(-\beta X_j)}$$

MS3:

$$B_{Y-\{\rho_3\}}(\rho_1) = B_{\{\rho_1,\rho_2\}}(\rho_1) = \frac{\exp(\beta X_1)}{\sum_{j=1,2} \exp(\beta X_j)}$$

MS4:

$$B_{Y-\{\rho_2\}}(\rho_1) = B_{\{\rho_1,\rho_3\}}(\rho_1) = \frac{\exp(\beta X_1)}{\sum_{j=1,3} \exp(\beta X_j)}$$

Importantly, each of the four choice set specifications is an assumed information processing strategy, aligned to the growing literature on attribute processing, which includes

---

[4] We thank a referee for suggesting this representation which we have modified to align with the four MSs.

[5] In ongoing research we are accounting for an additional endogenous response associated with each alternative in each choice set that identifies the acceptability of a specific alternative (Hensher and Ho 2014). Rose et al. (2014) also focus on joint estimation of choice and acceptability under RUM in the context of automobile type choice.

[6] We only consider model specifications that explicitly involve ρ_1 as best, but acknowledge that other model specifications could be considered such as $W_Y(\rho_3)$ *and* $W_{Y-\{\rho_3\}}(\rho_2)$.

consideration of relevant alternatives (see Hensher 2010 for an overview, also Rieskamp et al. 2006; and Leong and Hensher 2012a, b; Hensher and Ho 2014). Hence, the task is to establish under the specific treatment of alternatives (MS1–4), how a choice as a preference ordering (and resulting predicted choice probability ordering) is affected by adding (e.g., MS1) or removing (e.g., MS3) an alternative.[7] Critically, the alternatives deemed relevant can vary between choice tasks for an individual as a consequence of the combination of attribute levels describing each alternative, or for other less clear (possibly idiosyncratic) reasons.[8] Separate (mixed logit) models corresponding to different MSs are estimated to obtain not only the overall goodness of fit of each model (as one basis of determining the 'best' model in statistical terms) but also the predicted choice probabilities for each alternative under consideration.[9]

In recent years there has been growing interest within the discrete choice framework on asking respondents to select both the best and worst options from a set of alternatives. This literature recognises the additional behavioural information in the best and worst response mechanism (e.g., Flynn et al. 2007; Marley and Pihlens 2012; Collins and Rose 2011; Vermeulen et al. 2010). It is argued by various authors (see Louviere et al. 2013) that best–worst scaling delivers more efficient and richer discrete-choice elicitation than other approaches.

It is expected that the predicted choice probability distribution associated with MS3 will display greater similarity in the pairwise choice probabilities, and thus the model corresponding to MS3 will increase the risk of non-compliance (i.e., a reversal of preference ranking when the worst alternative is removed from the choice set). In contrast, the choice probability 'gap' is expected to be greater when the middle alternative is removed, a priori, in the MS4, which will increase the chance of compliance. Where the exploded best–worst formulation fits within the spectrum is unclear; however we speculate (and test below) that it is inferior to both binary forms[10] and the model for the selection of the best alternative in the full choice set (MS1).

The empirical setting

To investigate the compliance of the four MSs, we use a recent data set collected in Sydney that focussed on investigating preferences for a number of alternative road pricing reform packages for car users. Respondents were shown three alternatives and asked to indicate

---

[7] Including the implications we can draw in respect of ISDA as set out in Appendix 2.

[8] This may seem unclear for MS4 given that ISDA (Appendix 2) states that the rank ordering of all alternatives will not be changed if we remove from or add to the choice set an alternative which is statewise dominated by 'every state of the nature' (or every choice task in our application). The evaluation of the performance of MS4 where we actually remove the second best alternative, not the worst alternative, is best seen relative to MS3, but in both cases, the choice set on offer across the sample is 3 alternatives. It is the processing strategy that relates to the determination of the relevant set of alternatives. A dominated alternative can be any alternative in an uncertain preference ordered set.

[9] In the concluding section we discuss a way forward to jointly estimate a process decision rule model of the latent class form in order to reveal membership of each respondent to each class of MS, up to a probability.

[10] Note that the binary form accommodates differing mixes of two alternatives across the sample. This is not to be confused with a more traditional binary choice where the two alternatives are fixed across a sample.

which one was the best, and which one was the worst.[11] Given these two responses and a choice task of size three, we are able to identify the full preference ranking of the alternatives. The survey instrument was an online computer assisted personal interview accessed via laptops. Interviewers sat with the respondents to provide any advice that was required in working through the survey, while not offering answers to any of the questions. The data has been used in other papers (e.g., Hensher et al. 2013a, b) but not with the current focus.

The choice experiment consisted of three alternatives: the status quo and two labelled alternatives. The two labelled alternatives represent a cordon-based charging scheme and a distance-based charging scheme, randomly assigned to road pricing schemes 1 and 2. An illustrative choice screen, together with the boundaries of the proposed cordon-based charge area, is presented in Fig. 1. Each alternative was described by attributes representing the average amount of tolls and fuel outlaid weekly, the annual vehicle registration charge, and the allocation of revenues raised to improve public transport, improve and expand upon the existing road network, to reduce income tax, to contribute to general government revenue and to be used to compensate toll road companies for loss of toll revenue. The cordon-based charging scheme and a distance-based alternative were also described by a peak and off peak charge. Both non-status quo alternatives were also described by the year proposed that the scheme would commence.

A Bayesian D-efficient experimental design was implemented for the study. The design was generated in such a way that the cost related attribute levels for the status quo were first acquired from respondents during preliminary questions in the survey, whilst associated attributes for the cordon-based and distance-based charging schemes were pivoted off of these as minus percentage shifts representing a reduction in such costs for these schemes. Pivoted attributes included average fuel costs and annual registration fees. Fuel costs were reduced by anywhere between zero per cent and 50 per cent of the respondent reported values, either representing no reduction in fuel tax or up to a potential 100 per cent reduction in fuel taxes. Registration fees were reduced to between zero per cent and 100 per cent from the respondent-reported values (see Rose et al. 2008 for a description of pivot type designs). The toll cost was only included in the status quo alternative, being set to zero for the non status quo alternatives since it is replaced by the road pricing regime.

The allocation of revenues raised were fixed for the status quo alternative, but varied in the cordon-based and distance-based charging schemes over choice tasks. The allocation of revenue was varied from zero per cent to 100 per cent for a given revenue stream category. Within a charging scheme, the allocation of revenue was such that the sum had to equal 100 per cent across all possible revenue allocations.

The peak cordon charge varied between $2 and $20, whilst the off-peak cordon charge was varied between $0 and $15. The per kilometre distance-based charge for the peak period ranged from $0.05 to $0.40 per kilometre whilst the off-peak distance-based charge varied between $0 and $0.30 per kilometre. The ranges selected were based on ranges that we believe would contain the most likely levels if implemented (i.e., all reasonable 'states of nature'). The design was generated in such a way that the peak charges were always equal to or greater than the associated off peak charges. Finally, the cordon-based and

---

[11] The actual choice experiment had a fixed structure for the two choice questions, although both questions were shown at the same time. A referee pointed out that even if the participants were not forced to answer the questions in the order best, then worst, the order of the questions on the page might lead to that order for many participants. By showing both simultaneously we are unable to test for any sequence bias that others such as Hawkins et al. (2014) (In Press) did for a study where participants were free to make a best–worst pair of choices in either order.

Fig. 1 An illustrative choice screen and the location of the cordon-charge area

distance-based charging schemes were described by the year the scheme would be implemented. In each case, this was varied between 2013 (representing one year from the survey) and 2016 (representing a 4-year delay from the time of the survey).

The attributes and the relevant attribute levels for all alternatives are shown in Table 1. Priors for the design of the choice experiment were obtained from a pilot study of nine respondents collected prior to the main field phase. The final design consisted of 60 choice tasks which were blocked into 15 blocks. The blocking was accomplished by using an algorithm designed to minimise the maximum absolute correlation between the design attributes and the blocking column. Each respondent was shown sequentially four choice scenarios (60/15 = 4) each of three alternatives and asked to indicate their preferences in terms of best and worst scheme. This enables us to create a full rank order of the three alternatives.

The evidence

Mixed logit models (summarised in Appendix 1) were estimated in which the parameters associated with road pricing attributes were defined as random parameters with constrained triangular distributions. The overall goodness of fit as the log-likelihood at convergence

Table 1 The Choice experiment attribute levels and range

Attribute	Status quo	Cordon-based scheme	Distance-based scheme
Year scheme introduced	–	2013, 2014, 2015, 2016	2013, 2014, 2015, 2016
Average fuel per week	User reported level	0, −10, −20, −30, −40, −50 %	0, −10, −20, −30, −40, −50 %
Average toll per week	User reported level	$0.00	$0.00
Annual vehicle registration	User reported level	0, −25, −50, −75, −100 %	0, −25, −50, −75, −100 %
Peak cordon-based charge (per trip)	$0.00	$2.00, $6.50, $11.00, $15.50, $20.00	–
Off peak cordon-based charge (per trip)	$0.00	$0.00, $3.00, $6.00, $9.00, $12.00, $15.00	–
Peak distance-based charge (per km)	$0.00	–	$0.05, $0.12, $0.19, $0.26, $0.33, $0.40
Off peak distance-based charge (per km)	$0.00	–	$0.00, $0.06, $0.12, $0.18, $0.24, $0.30
% of funds allocated to public transport	0 %	0, 20, 40, 60, 80, 100 %	0, 20, 40, 60, 80, 100 %
% of funds allocated to road infrastructure	30 %	0, 20, 40, 60, 80, 100 %	0, 20, 40, 60, 80, 100 %
% of funds allocated to reducing tax	0 %	0, 20, 40, 60, 80, 100 %	0, 20, 40, 60, 80, 100 %
% of funds allocated to general revenue	65 %	0, 20, 40, 60, 80, 100 %	0, 20, 40, 60, 80, 100 %
% of funds allocated to private (toll) firms	5 %	0, 20, 40, 60, 80, 100 %	0, 20, 40, 60, 80, 100 %

(with pseudo R^2 and Akaike information criteria (AIC) in brackets) for each of the four models is as follows:

- MS1: −648.77 (0.254, 1.667) for the selection of the best alternative in the full choice set.
- MS2: −1353.63 (0.525, 1.712) for the exploded best–worst model.
- MS3: −432.76 (0.496, 1.127) for the most and second most preferred alternatives, and
- MS4: −338.70 (0.606, 0.892) for the best (most preferred) and worst (least preferred) alternatives.

The evidence suggests that the simple two-alternative choice set form of the best and worst alternatives (MS4) has a statistically better fit over all of the other processing forms. While this might have an element of 'obviousness', it is nevertheless an important piece of reinforcing evidence. What is important is that the best and worst alternatives vary across the sample (indeed even between choice sets for each respondent), and hence a choice set studied for a sample might contain many alternatives of which two specific ones are what really matters to a sampled individual evaluating a stated choice set.

Turning to the incidence of compliance, or more generally to an unconditional preservation of preference ordering of alternatives, in our application with four choice tasks per respondent, full compliance would be satisfied if the alternative chosen as the most preferred in each of the four choice tasks has the highest predicted choice probability.[12] The evidence is summarised in Table 2. This table is a sample wide summary based on choice probabilities predicted by each of the four models corresponding to four MSs. If SP data is to be the basis of application work, then we want to see which MS delivers the 'best prediction' in terms of the alternative stated as the most preferred having the highest predicted choice probability. The comparison of one MS against another is undertaken using the compliance test to establish which MS preserves the order (of most preferred) when other alternative choice set formations are considered.

The empirical evidence suggests that the specification that satisfies the compliance test best (but not 100 % across the sample) is the MS4 binary best–worst form. However, the compliance rate of this MS is not perfect and for the full sample of 200 car users, the average compliance rate is 3.24 out of 4 choice tasks. The MS1 full rank model (2.55/4) is an improvement over MS2, the exploded best–worst model (2.36/4). As might be expected, the best–second best form (MS3) does not satisfy the compliance test to the same extent (2.90/4) as the binary best–worst (MS4). This is intuitively plausible because MS3 narrows the difference in the choice probabilities for whatever behavioural reason (e.g., similarity, attraction), with an expectation that there will be more cases in which the most preferred alternative is no longer the one with the greatest predicted choice probability (i.e., a higher propensity of preference order switching). What this suggests is that a binary form based on the extremes in the ranks is more likely to satisfy the compliance test than any other form assessed (as promoted in Gourville and Soman 2007).

The findings in Table 2 can also be represented in terms of the average predicted probability of the best alternative, summarised in Fig. 2. As expected, given the evidence in Table 2, the mean choice probability associated with the best (most preferred) alternative varies from a high of 0.730 for the MS4, to a low of 0.437 for the exploded best–worst regime (MS2). The range, close to 0.30 probability units, is large and raises concerns about the behavioural validity of the exploded best–worst choice model form (noting

---

[12] Note that the alternative with the highest probability in each of the four evaluated choice tasks can be a different alternative.

Table 2 Average number of times (out of a max of 4) the alternative with highest, second highest and third highest probabilities is ranked first by each model specification

	Best alternative of the full choice set (MS1)	Exploded best-worst (MS2)	Best - Second Best (MS3)	Best - Worst MS4)
Highest probability	2.555	2.365	2.900	3.240
Second highest prob	0.985	1.050	1.100	0.760
Third highest prob	0.460	0.585	-	-
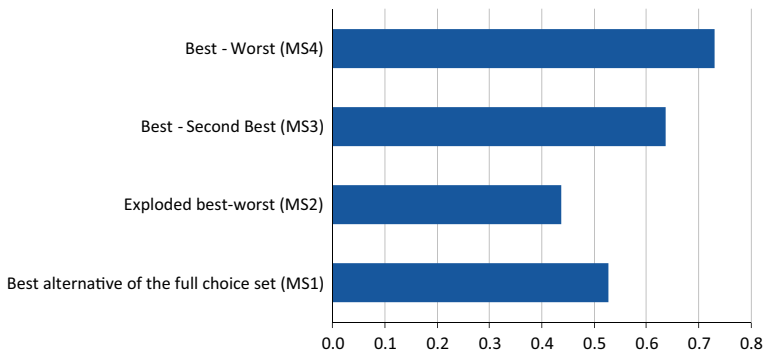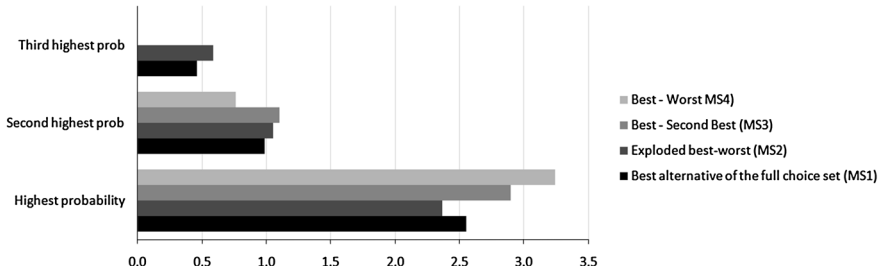
Fig. 2 Average predicted probability of the most preferred alternative by model specification

however that the evidence is based on only one data set). The second best regime is the best–second best, which has diluted part of the probability gains associated with the best in contrast to the worst alternative, as expected, since the best and second best alternatives are clearly closer in preference space (i.e., utility terms). What is of special interest is the evidence suggesting that the model for the selection of the best alternative in the full choice set (i.e. MS1) performs much better than the exploded best–worst regime (MS2).

A strong message from this evidence, if it can be replicated on a number of data sets, is that a simple binary form (with varying alternatives across the sample) aligns best with the compliance test and the consequent significant improvement in the ability of the choice model to predict a respondent's most preferred alternative. Whether there are some exogenous influences at work that might increase our understanding of the differences in the evidence on incidence of compliance is investigated in the following section.

Before investigating sources of influence on the incidence of compliance, a summary of an informative behavioural output, namely mean estimates of direct and cross elasticities

Table 3 Direct and cross elasticities under alternative choice set specifications

	Best of full choice set (MS1)	Exploded best–worst (MS2)	Best–2nd best (MS3)	Best–worst (MS4)
Number of alternatives	3	3 and 2	2	2
Peak period cordon charge				
Direct elasticities				
SC Alt1, 2	−0.052	−0.015	−0.029	−0.021
Cross elasticities				
SQ, Alt1, 2	0.031	0.006	0.020	0.013
Off-peak period cordon charge				
Direct elasticities				
Alt 1,2	−0.018	−0.006	−0.002	−0.013
Cross elasticities				
SQ, Alt1, 2	0.012	0.004	0.001	0.009
Peak period distance charge				
Direct elasticities				
SC Alt1, 2	−0.223	−0.022	−0.034	−0.190
Cross elasticities				
SQ, Alt1, 2	0.048	0.013	0.011	0.040
Off-peak period distance charge				
Direct elasticities				
SC Alt1, 2	−0.180	−0.006	−0.039	−0.137
Cross elasticities				
SQ, Alt1, 2	0.037	0.004	0.019	0.021

SC Stated choice*; SQ* Status quo

for the road pricing reform attributes associated with random parameters, is presented in Table 3. Even though the estimates are very small (highly inelastic), the important point to note is that the mean estimates are numerically very different between the four choice processing regimes. The highest mean estimates for direct elasticities are for the model of the best alternative in the full choice set and the preferred choice regime of best vs. worst, although the peak period cordon-based charge has similar direct elasticities between MS3 and MS4. The main message here is not one of the specific numerical estimates, but that the mean estimates (which are commonly used by practitioners) are very different between the four alternative processing regimes. We would argue that compliance might be an appealing guide in selecting the preferred model, and hence the associated elasticity estimates.

Systematic factors linked to compliance

Having identified the extent of compliance with ISDA across the sample, it is of value to see if there are some contextual and respondent-specific characteristics, as well as choice experiment features, that have a systematic link to the incidence of compliance. To identify candidate influences, we ran a series of ordered logit models in which the dependent variable was the discrete compliance rate for each respondent. The rate took the values 0–4
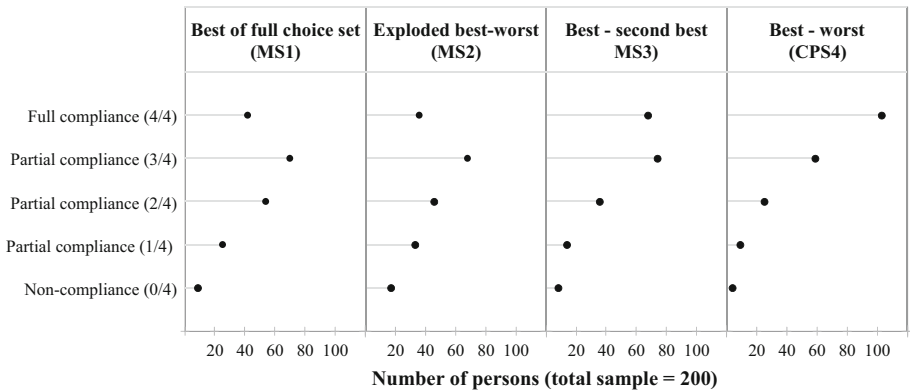
Fig. 3 Incidence of compliance by model specification

(see the frequency distribution given in Fig. 3). The results are summarised in Table 4, with partial (or marginal effects) available on request.

A number of variables were identified as statistically significant influences on the incidence of compliance. A particularly interesting result relates to the awareness of the road pricing debate. The positive parameter associated with awareness suggests that the incidence of compliance increases as the perceived awareness response increases. This highlights the role that increased knowledge of the context of the choice experiment plays in ensuring that the stated preference ordering is compliant.

Three socioeconomic characteristics (age, gender and personal income) are statistically significant in one, two or all of the four processing rule forms. As the age of the respondent increases we find, *ceteris paribus*, that the incidence of compliance increases under the exploded best–worst (MS2) specification but decreases under the best–second best choice processing rule (MS3). Personal income is only a statistically significant influence for the best–worst processing rule (MS4), with the negative sign. This suggests, *ceteris paribus*, that the incidence of compliance decreases as income increases. Gender has a negative influence in all models, suggesting that compliance is reduced for males compared to females. Overall, socioeconomic effects suggest that females on relatively low incomes tend to be associated with a preference ordering in choice experiments that complies better than other socioeconomic classes, with the age effect ambiguous.

The remaining six variables are trip-related attributes. Overall, with one exception (i.e., annual registration fee under best–worst model), the parameter estimates for all cost attributes are positive, suggesting, *ceteris paribus*, that as the cost of road use increases in the alternatives in the choice set, the compliance of the preference increases. This is, however, tempered when the road pricing reform regime involves a cordon-based charge in contrast to a distance-based charge.

Although these influences are statistically significant, they explain a very small amount of the variation in the incidence of compliance. Thus one might be tempted to conclude, given the available potential sources of systematic influence available in the data, that compliance of preference ordering is aligned strongly with individual-specific idiosyncratic effects which are not revealed.

Table 4 Sources of explanation of incidence of compliance: ordered logit (parameter estimates with t-values in brackets)

	Best of full choice set (MS1)	Exploded best–worst (MS2)	Best–2nd best (MS3)	Best–worst (MS4)
Number of alternatives	3	3 and 2	2	2
Sources of influence				
Awareness of road pricing debate (0–100)	0.3573 (3.25)	0.4885 (4.42)		0.6281 (5.80)
Age (years)		0.0066 (3.41)	−0.0037 (−1.96)	−0.0031 (−4.82)
Personal income ($000 s)	−0.3680 (5.76)	−0.2520 (−4.14)	−0.2287 (−3.76)	−0.522 (−7.41)
Gender (male = 1)		−0.0036 (−2.63)		0.0056 (3.35)
Distance travel per day (km)				−0.0003 (−4.43)
Annual registration fee ($)	0.0002 (2.25)	0.0036 (5.27)	0.0002 (2.45)	0.0132 (7.51)
Weekly fuel outlay ($)	0.0111 (9.82)	0.0063 (5.93)	0.0076 (7.57)	0.0175 (3.87)
Cordon-based peak period charge ($/day)	0.0105 (2.47)			
Distance-based charge, peak period ($/km)	0.0032 (1.93)			
Cordon charge scheme dummy (1,0)	−0.1770 (−2.57)		−0.1746 (−2.65)	−0.2677 (−3.69)
Constant	2.646 (36.4)	1.6920 (16.6)	3.1520 (28.2)	3.7640 (41.2)
Threshold parameters				
μ_1	1.4980 (27.2)	1.2950 (24.8)	1.0964 (17.2)	1.2464 (14.4)
μ_2	2.886 (53.4)	2.3420 (42.0)	2.3033 (39.8)	2.5144 (33.7)
μ_3	4.5060 (78.9)	3.9710 (65.9)	3.8890 (72.7)	3.9940 (57.5)
Log-likelihood at convergence	−5686.8	−6022.3	−5379.2	−4554.4
Log-likelihood with constants	−5792.7	−6081.3	−5432.9	−4718.4
AIC	2.849	3.016	2.694	2.283

In the mixed logit models reproduced in the Appendix, we have included a couple of socioeconomic effects and the reader will note that they are not statistically significant except for age in Model 4. We did investigate many more socio economic influences but we do find them to be statistically significant. Importantly, even though age is not statistically significant in the ordered logit model focussing on compliance, but it is in model 4, this cannot be interpreted as an inconsistency since the mixed logit model is designed to explore influences on choice which is not the same as influences that are systematically related to compliance

Conclusions

This paper has investigated a somewhat neglected issue in stated choice experiments; namely the relevance or otherwise of subsets of alternatives that are pre-specified in the design of choice experiments, given an individual's processing strategy. Although a number of authors such as Hensher (2006) have investigated the implications on attribute processing of the number of alternatives offered in a choice set and concluded that "As we increase the 'number of alternatives' to evaluate, *ceteris paribus*, the importance of considering more attributes increases, as a way of making it easier to differentiate between the alternatives.", the current paper takes a different approach to identifying the role of alternatives. That is, this paper focuses on the role that each alternative plays in establishing the preference ordering (and the *ex post* estimation associated choice probability) for an alternative.

Fundamentally, the question of interest is whether there is some redundancy in the offered set of 'statistically designed' alternatives[13] which may 'get in the way' of improving the capability of a model to predict the choice of the alternative that is stated as the most preferred (or best) alternative. The empirical analysis used subjective rank ordered preference data on road pricing reform alternatives obtained from a sample of respondents in a predefined stated choice experiment and predefined information processing rules.

Under the four choice making rules (or model specifications) considered, we have identified the extent to which compliance is satisfied. While all model specifications violate a full compliance condition *to some extent*, the binary best–worst processing strategy produces the most consistent rank ordering. Specifically, we find overwhelming empirical evidence to support a simple best–worst binary form, in contrast to other forms investigated; namely preservation of the (constructed) full rank choice set, the best versus second best, and the exploded best–worst form, the latter promoted in the broader literature of best–worst modelling. The evidence aligns with the contribution of Gourville and Soman (2007) who argue that respondents display an increased tendency to either of the extreme alternatives when the size of the choice set is increased (in our application, admittedly, of only up to three alternatives). Respondents are posited to increasingly rely on an all-or-nothing strategy.

The approach proposed and empirically assessed in this paper might be a strong candidate for establishing the most appropriate processing rule in choice making in future choice studies which can, *ex post*, be used to jointly estimate process and outcome choices. We encourage further testing using other data sets, especially where there are more than three alternatives offered in a choice experiment.

The findings are of immense relevance to transportation research, especially given the dominant role of SP or SC experiments which show in general a lack of consideration of process heuristics used by individuals when making choices amongst alternatives. In the existing literature, there is a strong (almost over) focus on the niceties of the statistical design of choice experiments (the first author can claim some amount of guilt on this from past contributions). If we are to continue to use SP experiments, then some awareness of process heuristics is really important, and if we can show repeatedly that a specific heuristic is a dominating one (e.g., best and worst only), then we can at least extract the set of

---

[13] Which are usually based on statistically respectable principles of experimental design, which contrast with the way that individuals actually process information in real or hypothetical situations.

parameter estimates that align closely with this.[14] This is not, we acknowledge, the basis of establishing the relevant choice set to use in model application for prediction and forecasts since this remains a major research and practical challenge when moving from a sample used in modelling to the population as a whole; but it is a good start and throws out the challenge for ongoing research in establishing possible segments of the population that adopt specific processing strategies. Current practices are far from ideal, and hence the need to recognise the signals offered by the evidence herein should be taken on board in research focussed on defining choice sets for population based forecasting.

There is a growing interest in the topic of this paper.[15] Alberini ([2014](#)) has also investigated the gains in statistical efficiency of considering the first and second best alternatives versus the best and worst; and also investigated choices in the presence and absence of a status quo alternative. While that paper is of interest in its own right as a contribution to exploring the role of subsets of alternatives (and a contrast of a single choice vs. offering five choice sets), it differs from our study in that we begin with a full rank order data of three alternatives and focus on investigating inferred processing strategies as guided by a number of model specifications. Hawkins et al. ([2014](#)) is another paper with some amount of common focus. They examined best and worst choices using discrete choice tasks, where participants selected either the best option from a set, the worst option, or selected both the best and the worst option. They found that the task (best, worst, or best and worst) does not alter the preferences expressed for the best (respectively, the worst) option, and also observed that the choice probabilities were consistent with a single latent dimension; that is, options that were frequently selected as best were infrequently selected as worst, and vice versa, both within and between respondents. They concluded that the diverse types of best and worst choices studied can be conceived as opposing ends of a single continuous dimension rather than distinct latent entities.

In ongoing research, using a random regret mixed logit model, we are jointly estimating a probabilistic decision rule model that can account for the role of each of the four MSs to see if we can further support the dominance of a specific processing heuristic. The joint estimation approach can replace the separate models for each model specification as well as condition the probability allocation on the socioeconomic and other contextual influences. The ongoing joint estimation is in part an econometric refinement, but it provides an additional basis of establishing the role of each MS in choice set definition at the respondent level.

---

[14] The typical stated choice design uses algorithms that are focussed on criteria such as D-optimality, utility balance etc. that are unrelated to the behavioural perspective that ISDA (Appendix [2](#)) relates to. Essentially we might suggest that what is missing in the design of choice experiments is how one can accommodate the processing strategies of each sample respondent. This is a dynamic design process since one does not know in advance much about how a respondent might process a specific choice set. We have started looking at this and have some simple attribute non attendance (ANA) designs in NGene (Choice Metrics) but they are not fully tested. What we are doing in this paper is to *ex post* revisit the responses to the choice experiment and with the richness of rank order response data we can test a number of processing strategies. This is hopefully the beginning of a more serious interest in the challenge of how one might eventually embed process heuristics in choice experiments.

[15] These two papers were drawn to our attention by a referee.

Appendix 1

See Appendix Table 5.

Table 5 Summary of Mixed Logit Models (t-values in brackets)

	Model 1 (MS1)	Model 2 (MS2)	Model 3 (MS3)	Model 4 (MS4)
Random parameters				
Cordon-based charging, peak period ($/week)	−0.0568 (−3.02)	−0.0163 (−1.92)	−0.0514 (−2.54)	−0.0541 (−2.73)
Cordon-based charging, off-peak period ($/week)	−0.0168 (−1.15)	−0.0066 (−0.78)	−0.0027 (−0.19)	−0.0299 (−1.55)
Distance-based charging, peak period ($/week)	−0.0509 −4.41)	−0.0054 (−1.99)	−0.0224 (−2.43)	−0.0678 (−4.52)
Distance-based charging, off-peak period ($/week)	−0.0409 (−3.85)	−0.0017 (−0.73)	−0.0285 (−3.13)	−0.0434 (−3.31)
Non-random parameters				
Annual registration charge ($)	−0.0019 (−7.71)	−0.0005 (−3.63)	−0.0015 (−5.22)	−0.0021 (−6.40)
Fuel cost ($/week)	−0.0241 (−3.86)	−0.0182 (−4.62)	−0.0237 (−3.40)	−0.0280 (−2.89)
Toll outlay ($/week)	−0.0655 (−4.18)	−0.0016 (−0.39)	−0.0766 (−4.14)	−0.0759 (−3.65)
New charges introduced after 2014 (1, 0)	0.2711 (1.90)	0.1550 (1.59)	0.2371 (1.48)	0.3053 (1.47)
Cordon-based charge dummy (1, 0)	−2.077 (−3.94)	−1.516 (−4.51)	−2.387 (−3.94)	−1.4473 (−2.03)
Distance-based charge dummy (1, 0)	−2.096 (−3.81)	−1.719 (−4.98)	−2.4199 (−3.86)	−1.5829 (−2.12)
Number of weekly trips that passed through CBD	−0.0287 (−0.69)	−0.0574 (0.34)	−0.0694 (−1.58)	−0.0160 (−0.27)
Support road pricing trial (1, 0)	1.812 (5.89)	1.2603 (5.49)	1.8259 (4.84)	1.9263 (4.43)
Age (years)	−0.0097 (−1.27)	0.0023 (0,55)	−0.0022 (−.26)	−0.0255 (−2.53)
Personal income ($000 s)	0.0004 (0.19)	0.0024 (0.58)	0.0014 (−.57)	0.0017 (0.67)
Per cent of revenue allocated to				
Improving public transport	0.0162 (5.40)	0.0074 (3.91)	0.0163 (4.62)	0.0163 (3.99)
Improving existing and construct new roads	0.0104 (3.77)	0.0031 (1.69)	0.0089 (2.79)	0.1226 (3.27)
Reducing personal income tax	0.0118 (3.90)	0.0062 (3.19)	0.0135 (3.76)	0.0105 (2.62)
Private toll road companies	0.0034 (1.03)	−0.0004 (−0.17)	0.0054 (1.40)	0.0003 (0.07)
Model fit				
Log-likelihood (0)	−878.9	−2866.8	−878.9	−878.9
Log-likelihood at convergence	−648.7	−1353.6	−432.8	−338.7
Adjusted pseudo R^2	0.254	0.525	0.496	0.606
AIC (sample adjusted)	1.667	1.712	1.127	0.892

Model 2 is quite different to the other models in that it includes the additional choice set based on the worst alternative excluding the first best

Appendix 2: A theoretical setting

A theoretical basis aligned with the identification of compliance with the highest predicted choice probability associated with the most preferred alternative is the Axiom of

Irrelevance of Statewise Dominated Alternatives (ISDA) proposed by Quiggin (1994). Although it is set out in the context of a lottery (i.e., a single attribute), it can easily be applied in a context of more than one attribute associated with an alternative. Quiggin's (1982, 1994) principle of ISDA states that a choice (as a preference ordering) from any given choice set is not affected by adding *or* removing an alternative that is inferior for 'every state of the world' (also known as a statewise context described by alternatives, choice sets and underlying attributes of alternatives). Inferiority can be aligned to estimates of relative utility (or choice probability) associated with alternatives in discrete choice models that assume each agent acts as if they are a utility maximiser.[16] The satisfaction (or not) of this principle is also related to the general study of context effects in choice (elegantly presented in, for example, Rieskamp et al. 2006). Formally, alternative A is statewise dominant over alternative B if A gives a better outcome than B in every possible future state (with the assumption that the chooser knows what attributes associated with each alternative matter).[17] The set of possible future states is defined in our situation as offered choice sets, which are assumed to reasonably establish the domain of all relevant attribute packages or prospects associated with an alternative.

Statewise dominance is aligned with stochastic dominance as a form of stochastic ordering in the presence of risk and uncertainty. Also known as state-by-state dominance, it is one of the simplest cases of stochastic dominance, defined as follows: uncertain prospect A is statewise dominant over uncertain prospect B if A gives a better outcome than B in every possible future state.[18] Specifically, the term refers to situations where an uncertain outcome (as a probability distribution over possible outcomes, also known as prospects) can be ranked as superior to another uncertain outcome, based on preferences regarding outcomes. A preference might be a simple ranking of alternatives from most preferred to least preferred, as is common in the rank ordering of alternatives in choice sets related to SP experiments. Formally, a real random variable A is less than a random variable B in the "usual stochastic order" if $\Pr(A > x) \leq \Pr(B > x)$ for all $x \, \varepsilon(-\infty, \infty)$, where $\Pr(.)$ denotes the probability of an event. If additionally $\Pr(A > x) < \Pr(B > x)$ for some x, then A is stochastically strictly less than B. For example, if a road pricing reform package A has a lower relative disutility than another reform package B (accounting for the utility weights that an individual assigns to each attribute), then A gives a better outcome than B (in respect to the choice probability distribution; hence allowance for uncertainty).[19] Anyone who prefers more to less (i.e., who has monotonically increasing preferences) will always prefer a statewise dominant prospect.

Quiggin (1994) comments that "the most obvious danger of a preference revelation procedure dependent on the set of alternatives is that manipulation[20] of the set of the alternatives may yield irrational choices" (p. 156). There are many ways to manipulate the choice set of alternatives and introduce statewise dominated alternatives into the choice set (see Dequiedt and Martimort 2006). The question of particular interest is how can we test whether a particular manipulation actually yields irrational choices? The simplest form of

---

[16] Quiggin was interested in regret theory under uncertainty rather than utility maximising theory, but the logic is the same.

[17] In contrast to the analyst or modeller who only observes preferences and choices up to a probability due to the inability to identify all sources of utility that matter to a decision maker.

[18] More precisely, at least as good an outcome in every state, with strict inequality in at least one state.

[19] Examples are often couched in lottery terms. For example, if a dollar is added to one or more prizes in a lottery, the new lottery statewise dominates the old one.

[20] In stated choice experiments by the way it is designed.

manipulation arises if the introduction of relatively unattractive alternatives (or very similar, but slightly less attractive alternatives) influences decisions over the remaining choices, in respect of preference order. Stated choice experiments that are designed on statistical properties in contrast to behavioural conditions may well fall victim to this situation as the data context used to reveal preferences.

In the context of this paper, we are interested in the extent to which any pairwise choice (where say A is preferred to B) is likely to be manipulated (in the sense of the resultant choice probability order) by the addition of a less attractive alternative C, to the choice set. This less attractive alternative may be offered up in the statistical design of a choice experiment.[21] If the addition of C results in a switch of preferences between A and B in terms of rank order, a respondent can extract additional relative utility (and hence the choice model reflects this with adjustments in the choice probabilities). Certainly, under whatever processing strategies (represented as a modelling specification in this paper) the respondent may use, the addition or exclusion of a specific alternative will always result in a distribution of choice probabilities predicted by discrete choice models; however the issue of greatest concern is whether a specific processing strategy results in a re-ordering of the preferences (as well as magnitude adjustments for a given order) as reflected through the modelling paradigm.

To set out the fundamental ISDA axiom of Quiggin, in the context of a multiattribute choice setting of current interest,[22] we need to introduce some notation. Before doing so, we need to explain the link between Quiggin's application involving a single attribute (a lottery price) with a given set of outcomes, which enables a direct assessment of dominance, and the focus of our study. When we move to a multiattribute setting (see also footnote 31), the notion of dominance is based on utility maximising behaviour which assumes that multiple attributes can be converted into a single measure (i.e., indirect utility 'index') for comparison. Since we cannot identify dominance through an attribute-by-attribute comparison, we have to redefine it through the specification (and estimation) of the utility expression for each alternative which maps into the ranked order. This ensures an interpretation of preference ordering in which we identify the choice probabilities associated with each of the ranked alternatives (given the model specification). Given the assumption of a mapping between stochastic dominance and ISDA, we are able to use this approach under a multiattribute specification to test compliance with ISDA, which contrasts with what one would do with a simply lottery setting when one has a lottery price and an associated financial prize and associated winning odds.

---

[21] This may be quite different to the set of alternatives that is the respondent's consideration set in a real market.

[22] What we are doing here is moving beyond the somewhat simple setting of a single price attribute attached to a lottery (which is a common example in experimental economics) and the one discussed by John Quiggin, to a situation common in transport studies where there is more than one attribute at play in choice making. The multiattribute setting is essentially a utility-based evaluation of alternatives described by a series of attributes weighted by their marginal (dis)utility contribution. An alternative is said to be state-wise dominated if it always gives a lower marginal utility than another alternative in all states of the world (i.e., choice tasks) for a given specification of the utility function. Assuming the utility maximisation behaviour of decision makers, a state-wise dominated alternative can be identified from the observed data as any alternative that is not ranked best by a respondent. The behavioural implications as we see it are exactly the same except that more than one attribute is being processed. We faced this same matter in our earlier work on extending expected utility theory to incorporate risk attitude in the context of measuring the value of expected travel time savings (Hensher et al. 2011) where there were a number of attributes, noting that the dominant literature in economics had focussed on the gamble (or lottery) with one attribute and indeed did not even assess the contribution of that attribute to choice by including a marginal disutility or utility parameter estimate (presumably because in a single attribute setting it was not necessary!).

In Quiggin's notation (slightly modified to fit our application), consequences arise from the interaction between individual choices and the occurrence of state of the world n, where n is one of a set of N possible states (or choice sets) of the world, each defined by a set of observed and unobserved attributes associated with each alternative in the choice set, occurring with probability π_n. Under a binary choice situation, if an individual chooses action A_i in preference to A_j, given that the nth state of the world occurs (i.e., this specific choice set of alternative outcomes is all that is offered), the actual consequence is x_{in}; however had the individual chosen differently, x_{jn} would have occurred. It is assumed that the x_n are members of a set X of consequences (such as x_{in}, x_{jn}), ordered by a preference relationship. Denoting the level of utility arising from this choice as $v(x_{in}, x_{jn})$, we define $v^*(x_{in}, x_{jn}) = v(x_{in}, x_{jn}) - v(x_{jn}, x_{in})$, as the net gain, in utility terms, of choosing A_i rather than A_j in the event that state n occurs. Under utility maximisation, it is well known that for a binary preference revelation setting, if one prospect delivers a better outcome in every state (or choice set), it will be preferred, preserving statewise stochastic dominance.

We cannot ensure that this condition holds when there are more than two alternatives. What are the conditions applying to the function v^* that will ensure that the introduction of statewise dominated alternatives (be they the second or third ranked in a 3-alternative choice set across all choice sets representing the state of nature) will not affect the ranking of any other pair of actions? A proposed resolution is Quiggin's ISDA Axiom.

ISDA is defined as follows: Let A_i be the most preferred alternative of the set of alternatives A offered in a stated choice experiment, and suppose there are some alternatives A_r, A_s ε A such that A_r statewise dominates A_s. Then A_i is the most preferred alternative of the set of alternatives $A^* = A - \{A_s\}$. Quiggin shows (for three alternatives) that this requirement is sufficient to give a complete characterisation of the properties of v. Define $v^*(x_{in}, x_{jn}, \{x_{kn}\}) = v(x_{in}, \{x_{kn}\}) - v(x_{jn}, \{x_{kn}\})$. $\{x_{kn}\}$ is very general notation to represent a finite set of alternatives in which $k \neq i$, that allows us to recognise the finite number of alternatives which is greater than a binary set. When expressed as $v(x_{in}, \{x_{kn}: k \neq i\})$, for example, the object is to maximise the expectation of v. The expectation of v^*, $E[v^*]$ is positive *iff* A_i is preferred to A_j given the set of alternatives A.

Assume that ISDA holds. Then v^* may be represented by a function, say g, of the form $v^*(x_{in}, x_{jn}, \{x_{kn}\}) = g(x_{in}, x_{jn}, \max(\{x_{kn}\}))$. For three alternatives, a_1, a_2, a_3, $v^*(a_1, a_2, \{a_1, a_2, a_3\})$ can depend on a_3 only if $a_3 > a_1$, $a_3 > a_2$; that is, if $a_3 = \max\{a_1, a_2, a_3\}$. Hence v^* may be represented by a function of the form $g(.)$ for our purposes of implementing ISDA.[23] This is the basis of the empirical test undertaken in this paper on a number of model specifications of the choice model. Specifically, we have to identify the choice probability order under a number of processing strategies in choice making, and establish the incidence of compliance with ISDA. This then provides one useful reference point on which to select an application choice set. We are not aware of any (theoretical) test having been proposed and implemented in the discrete (stated) choice literature to guide the identification of choice set formation (in terms of 'relevant' alternatives) that aligns with particular model specifications as presented in the main text.

In conclusion, we recognise that one needs to know "all future states of the world" ahead of time, and in the context of SC experiments this is what we are assuming, as is a common setting for such experiments. It may be a criticism of such experiments in that alternatives in choice sets are 'imposed' on respondents who are asked to assess the offered alternatives and typically choose the most preferred. However, given this setting, the ISDA criterion has much merit and can be used herein as a theoretical setting to assess four

---

[23] Quiggin (1994) presents a one page proof of this proposition (which he calls Proposition 1).

reasonable heuristics about how a respondent processes alternatives that they have been asked to assess, as defined by the statistical design constructed by the analyst as if all alternatives are relevant.

References

Alberini, A.: Repeated questioning in choice experiments: are we improving statistical efficiency or getting respondents confused? J. Environ. Econ. Policy 1(2), 216–233 (2014). doi:10.1080/21606544.2012.698516

Collins, A.T., Rose, J.M.: Estimation of stochastic scale with best–worst data. Working paper, Institute of Transport and Logistics Studies, The University of Sydney Business School (2011)

Dequiedt, V., Martimort, D.: Non-manipulable mechanisms: a brief overview. Public Econ. 18–19, 3–16 (2006). http://economiepublique.revues.org/4392

Flynn, T.N., Louviere, J.J., Peters, T.J., Coast, J.: Best–worst scaling: what it can do for health care research and how to do it. J. Health Econ. 26, 171–189 (2007)

Gourville, J.T., Soman D.: Extremeness seeking: when and why consumers prefer the extreme, Harvard Business School Marketing Research Working Paper Number 07-092 (2007)

Hawkins, G.E., Marley, A.A.J., Heathcote, A., Flynn, T.N., Louviere, J.J., Brown, S.D.: The best of times and the worst of times are interchangeable. Decision 1(3), 192–214 (2014). doi:10.1037/dec0000012

Hensher, D.A.: How do respondents process stated choice experiments?—Attribute consideration under varying information load. J. Appl. Econ. 21(6), 861–878 (2006)

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

Hensher, D.A., Ho, C.: The role of perceptual acceptability of alternatives in identifying and assessing choice set processing strategies under random regret minimisation, November (2014)

Hensher, D.A., Rose, J.M., Collins, A.: Understanding buy in for risky prospects: incorporating degree of belief into the *ex ante* assessment of support for alternative road pricing schemes. J. Transp. Econ. Policy 47(3), 453–473 (2013a)

Hensher, D.A., Collins, A.T., Greene, W.H.: Accounting for attribute non-attendance and common-metric aggregation in a latent class mixed multinomial logit model: a warning on potential confoundment. Transportation 40(5), 1003–1020 (2013b)

Hensher, D.A., Greene, W.H., Li, Z.: Embedding risk attitude and decisions weights in non-linear logit to accommodate time variability in the value of expected travel time savings. Transp. Res. Part B 45, 954–972 (2011)

Leong, W., Hensher, D.A.: Embedding decision heuristics in discrete choice models: a review. Transp. Rev. 32, 313–331 (2012a)

Leong, W., Hensher, D.A.: Embedding multiple heuristics into choice models: an alternative approach. J. Choice Model. 5, 131–144 (2012b)

Leong, W., Hensher, D.A.: Contrasts of relative advantage maximisation with random utility maximisation and regret minimisation. J. Transp. Econ. Policy (in press 2014). Accepted 4 Dec 2013

Louviere, J.J., Lings, I., Islam, T., Gudergan, S., Flynn, T.: An introduction to the application of (case 1) best–worst scaling in marketing research. Int. J. Mark. Res. 30, 292–303 (2013)

Marley, A.A.J., Flynn, T.N.: Best and worst scaling: theory and application. In: James D. Wright (Editor-in-Chief), International Encyclopedia of the Social and Behavioral Sciences, 2nd Edition. Elsevier (2015)

Marley, A.A.J., Pihlens, D.: Models of best–worst choice and ranking among multi-attribute options (profiles). J. Math. Psychol. 56, 24–34 (2012)

Quiggin, J.: A theory of anticipated utility. J. Econ. Behav. Organ. 3(4), 323–343 (1982)

Quiggin, J.: Regret theory with general choice sets. Risk Uncertain. 8(2), 153–165 (1994)

Rieskamp, J., Busemeyer, J., Mellers, B.: Extending the bounds of rationality: evidence and theories of preferential choice. J. Econ. Lit. 44, 631–661 (2006)

Rooderkerk, R.P., Van Heerde, H.J., Bijmolt, T.H.A.: Incorporating context effects into a choice model. J. Mark. Res. 48(August), 767–780 (2011)

Rose J.M.: Interpreting discrete choice models based on best–worst data: a matter of framing, 93rd Annual Meeting of the Transportation Research Board TRB 2014, Washington D.C., United States, 16th January 2014

Rose, J.M., Bliemer, M.C., Hensher, D.A., Collins, A.T.: Designing efficient stated choice experiments in the presence of reference alternatives. Transp. Res. Part B 42(4), 395–406 (2008)

Rose, J.M., Beck, M.J., Hensher D.A.: Eliminating endogeneity bias in the calibration of choice experi-
ments: the joint estimation of choice and choice certainty and choice and alternative acceptability,
Submitted to Transportation Research Part A, being revised, November (2014)

Trueblood, J.S., Brown, S.D., Heathcote, A., Busemeyer, J.R.: Not just for consumers: context effects are
fundamental to decision–making. Psychol. Sci. **24**, 901–908 (2013)

Vermeulen, B., Goos, P., Vandebroek, M.: Obtaining more information from conjoint experiments by best–
worst choices. Comput. Stat. Data Anal. **54**, 1426–1433 (2010)

David Hensher is Professor of Management, and Founding Director of the Institute of Transport and
Logistics Studies (ITLS) at The University of Sydney. A Fellow of the Australian Academy of Social
Sciences, Recipient of the 2009 International Association of Travel Behaviour Research (IATBR) Lifetime
Achievement Award in recognition for his long-standing and exceptional contribution to IATBR as well as
to the wider travel behaviour community; Recipient of the 2006 Engineers Australia Transport Medal for
lifelong contribution to transportation. David is also the recipient of the Smart 2013 Premier Award for
Excellence in Supply Chain Management. Honorary Fellow of Singapore Land Transport Authority, and a
Past President of the International Association of Travel Behaviour Research. He has published extensively
(over 550 papers) in the leading international transport journals and key journals in economics as well as 12
books. David has advised numerous government and industry agencies in many countries (notably Australia,
New Zealand, UK, USA and The Netherlands), with a recent appointment to Infrastructure Australia's
reference panel on public transport, and is called upon regularly by the media for commentary.

Chinh Ho is a Senior Research Fellow in the Institute of Transport and Logistics Studies (ITLS) at The
University of Sydney. He completed his Ph.D. in 2013 on intra-household interactions and group decisions
in travel mode choice, with a particular focus on planning strategies and segments of car users to encourage
the use of more sustainable travel modes. Chinh was awarded the John H Taplin Prize for the best paper
presented at the Australasian Transport Research Forum 2012 and the Innovation Grant from World
Conference on Transport Research 2013. Chinh also holds a Bachelor of Civil Engineering with first class
Honours in Bridge and Highway Engineering from Ho Chi Minh City University of Technology, Vietnam
where he was a full-time lecturer. Before joining ITLS Chinh obtained a masters degree in Transport
Planning from Nagoya University, Japan.