WORKING PAPER

ITLS-WP-14-09

Recognising the irrelevance of statewise–dominated alternatives in defining the composition of a choice set

By
David A. Hensher and Chinh Ho

May 2014

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in Transport and Logistics Management

The University of Sydney

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

NUMBER: Working Paper ITLS-WP-14-09

TITLE: **Recognising the irrelevance of statewise–dominated alternatives in defining the composition of a choice set**

ABSTRACT: This paper is motivated by the primary idea (or curiosity) that the distribution of choice probabilities associated with a set of alternatives defining a given choice set provides strong evidence on the way that agents appear to process the description of each alternative in a stated choice experiment, conditional on other contextual influences that are agent specific. The supplementary interest is in the extent to which the established probability distribution, given the ranking of a set of alternatives, is able to be the basis of establishing whether a specific decision rule (within a utility maximising setting) offers the preferred behavioural 'explanation' of which alternatives really matter in choice making. Examples of interest include decision rules such as the relevance, in a rank order of alternatives, of all offered alternatives, variants of best-worst, and first best-second best. The underlying theoretical context to guide the preference ruling (or candidate alternatives under a rank order) is the Axiom of Irrelevance of Statewise Dominated Alternatives (ISDA) proposed by Quiggin (1995). In this paper we use a choice experiment on road pricing reform scenarios to illustrate a way to determine, under utility maximisation and knowledge of the full rank order of offered alternatives in a choice experiment, which set of alternatives satisfies ISDA and hence is a preferred choice set to use in estimation and application of a choice model.

KEY WORDS: ***best-worst, first best, second best, choice models, choice sets, full rank, irrelevance of statewise dominated alternatives, road pricing reform, choice experiment***

AUTHORS: **Hensher and Ho**

Acknowledgements: The research contribution is linked to an Australian Research Council Grant DP140100909 (2014-2016) on 'Integrating Attribute Decision Heuristics into Travel Choice Models that accommodate Risk Attitude and Perceptual Conditioning'. Discussions on best-worst methods with Andrew Collins and John Rose are appreciated. Advice from John Quiggin is also much appreciated, especially his feedback that 'this seems like a really great use of the ISDA idea'.

CONTACT: INSTITUTE OF TRANSPORT AND LOGISTICS STUDIES (C13)

The Australian Key Centre in Transport and Logistics Management

The University of Sydney NSW 2006 Australia

Telephone: +612 9114 1824

E-mail: business.itlsinfo@sydney.edu.au

Internet: http://sydney.edu.au/business/itls

DATE: May 2014

1. Introduction

Stated choice experiments have proven popular as a framework within which to construct choice sets of alternatives, and to seek respondent preferences through a number of responses such as choosing the preferred alternative, rank ordering the full set of alternatives, or choosing the best and worst. Within the setting of utility maximisation, these various response metrics are all candidate derivatives of the full rank order.

It has become common practice to preserve all alternatives in model estimation, including the form where the best and worst are identified and a choice explosion method is used which defines the initial full choice set with rank=1 as the chosen alternative and then removes this alternative and defines the worst rank as the chosen (but reversing the attribute sign) (See Rose 2014). However, the selection of the number of alternatives in a choice set, as well as contrasts across choice sets of the same alternatives with coverage of the attribute space that is likely to represent all 'states of nature', 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 might be considered[1], we can attempt to gain a better behavioural understanding of the extent to which all alternatives are value adding in preference revelation, and whether a subset of the alternatives, differing across a sample of individuals, is all that really matters in developing a practical model to predict the choice of a preferred alternative. This raises many questions of behavioural curiosity, including whether the full set of alternatives should be preserved in estimation or whether some other paradigm of choice set specification might be more reflective of the role of each alternative.

The paper is structured as follows. We begin by presenting an overview of a theoretical perspective offered in the broader literature on choice and uncertainty, centred on the Axiom of Irrelevance of Statewise Dominated Alternatives (ISDA). This provides the basis of a way forward in testing the compliance under utility maximisation of a number of alternative ways of representing the processing strategies of respondents, given a known full rank order of alternatives.

We draw on a choice experiment designed to study car user preferences for alternative road pricing reform packages, as a way of empirically identifying the incidence of compliance with ISDA. The findings are then presented, obtained from estimation of a series of mixed logit models, followed by interpretation and an inquiry into factors that may explain in part the relative incidence of compliance with ISDA across the studied processing rules. The paper concludes with a summary of the main findings and what this means for ongoing discrete choice applications.

2. The Theoretical Setting

To guide us in identifying a way forward in determining the set of choice sets that pass muster in respect of consistent preference ordering, we need some theoretical basis, given that the model aligns with the behavioural choice rule that each agent acts as if

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

they are a utility maximiser. An appealing axiom to guide us is the Axiom of Irrelevance of Statewise Dominated Alternatives (ISDA) proposed by Quiggin (1995). Quiggin's (1982) 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'. Inferiority can be aligned to estimates of relative utility (or choice probability). Formally, alternative A is statewise dominant over alternative B if A gives a better outcome than B in every possible future state. This is defined in our situation as offered choice sets, which 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. 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.[2] 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. It is based on preferences regarding outcomes. A preference might be a simple ranking of outcomes from preferred to least preferred, as is common in the rank ordering of alternatives in choice sets related to stated choice 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 $x \in (-\infty,\infty)$, 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 has a lower relative disutility than another reform package (accounting for the utility weights that an individual assigns to each attribute), then the outcome is better[3]. Anyone who prefers more to less (i.e., who has monotonically increasing preferences) will always prefer a statewise dominant prospect.

Quiggin (1995) comments that the most obvious danger of a procedure dependent on the set of alternatives is that 'manipulation' (see Dequiedt and Martimort 2006) of the set of the alternatives may yield irrational choices. The question of particular interest is how can we test whether this situation actually occurs in a particular choice experiment? The simplest form of 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.

In the context of this paper, we are interested in the extent to which any pairwise choice (where say *A* is preferred to *B*) can be manipulated (in the sense of the resultant choice probability order) by the addition of a less attractive alternative *C*, to the choice set. If the addition of *C* results in a switch of preferences (specifically preference order) between A and *B*, the manipulator (i.e., the respondent) can extract additional relative utility (and hence the choice model translates this into adjustments in the choice probabilities). While the addition or exclusion of a specific alternative will always result in changes in the choice probabilities of each remaining alternative, the issue of greatest concern is whether such processing conditions result in a re-ordering of the preferences as reflected through the modelling paradigm. We begin with the view that a

[2] More precisely, at least as good an outcome in every state, with strict inequality in at least one state.
[3] 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.

respondent's stated rank order is reasonably assumed to be 'reliable' in choice experiments where non-strategic behaviour is the behavioural norm (in contrast to strategic voting ranks designed to manipulate, for example, public opinion), and hence we are interested in finding a way in which we can test for non-manipulation and preservation of irrational choices under the presence of alternatives additional to a preference structure under a pairwise choice.

To set out the fundamental ISDA axiom of Quiggin, we need to introduce some notation. In Quiggin's notation (slightly modified), 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 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 n^{th} state of the world occurs, 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, 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. The function v is non–decreasing in its first argument and non–increasing in its second, and this property carries over to v^*. 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 statewide 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, and suppose there are some alternatives $A_r, A_s \in \mathrm{A}$ such that A_r statewise dominates A_s. Then A_i is the most preferred alternative of the set of alternatives $\mathrm{A}^* = \mathrm{A} - \{A_s\}$. Quiggin shows (for three alternatives) that this requirement is sufficient to give a complete characterisation of the properties of v. Define $v^*(x_{in}, x_{jn}, \{x_{kn}\}) = v(x_{in}, \{x_{kn}\}) - v(x_{jn}, \{x_{kn}\})$. 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 g of the form $v^*(x_{in}, x_{jn}, \{x_{kn}\}) = g(x\text{in}, x\text{jn}, \max(\{x_{kn}\}))$. 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 $g(.)$. This is the basis of the empirical test undertaken below on a number of processing forms of the choice model. Specifically, we have to identify the choice probability order under a number of process rules 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 test having been proposed and implemented in the discrete choice literature to guide the identification of a behaviourally preferred choice set.

3. Alternative Processing Specifications of Choice Sets

Given a respondent's rank order of alternatives as a revelation of preference, in a three-alternative choice set there are a number of interesting outcome forms worthy of consideration as candidate processing strategies. There are (at least) four of particular interest, within the binary outcome metric set to 1 for the most preferred and 0 for the other alternatives: P1=preserving all alternatives in a 3-alternative choice set; P2= redefining the choice set as two choice sets, with the first the same as P1 and the second as the worst choice (removing the alternative that is best for each respondent and changing the sign on attributes associated with the remaining two alternatives); P3=including only the 1st and 2nd best alternatives; and P4=including the 1st best and the worst alternative for each respondent.

Importantly, each of the four specifications is an assumed information processing strategy, aligned to the growing literature on attribute processing, which includes consideration of relevant alternatives (see Hensher 2010 for an overview). As such, the determination of compliance with ISDA is within the context of a given processing strategy linked to how three alternatives (in our examples) are processed. Critically, the 'relevant' alternatives can vary between choice sets for an individual as a consequence of the offered attribute package associated with each offered alternative, or for other less clear (possibly idiosyncratic) reasons[4].

Separate (mixed logit) models (P1-P4) 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 choice probabilities estimated for each alternative under the four processing rules. These choice probabilities (or estimated outcomes across all states of nature – the latter being four choice sets per respondent), are the essential information used to identify the incidence of compliance with ISDA. This is our theoretical benchmark in determining which processing rule is a better representation in order to ensure that the 1st best retains its rank order in the choice probability outcome across all choice sets, as an assumed representation of all states of nature.

In recent years there has been growing interest within the discrete choice framework on seeking responses to scenarios where respondents select both the best option and the worst option 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). 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; however we are unaware of this preference for P2 being assessed within a setting, given a known rank order, that considers other processing (or elicitation) rules *against* a theoretical benchmark such as ISDA.

---

[4] This may seem unclear for P4 given that ISDA 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 P4 where we actually remove the second best alternative, not the worst alternative, is best seen relative to P3, 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.

In addition to the 'traditional best-worst' processing paradigm and the standard full choice set approach, we also consider two variants. One involves removing the worst ranked alternative and modelling the choice between the 1st and 2nd best alternatives; the other involves the two 'extreme' alternatives in preference space, modelling the choice between the 1st best and worst alternative from the full rank order. It is expected that the probability distribution associated with the 1st and 2nd best regime will display greater similarity in choice probabilities, and thus the former variant will increase the risk of violation of ISDA. In contrast, the choice probability 'gap' is expected to be greater, *a priori*, in the 1st best and worst regime, with a well specified model, which will increase the chance of ISDA compliance. Where the traditional best-worst formulation fits within the spectrum is unclear; however we speculate (and test below) that it is inferior to both binary forms[5] and the full rank order.

4. The Empirical Setting

To investigate the compliance of the four processing rules with ISDA, under utility maximisation, 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. The survey instrument was an online computer assisted personal interview accessed via laptops used by interviewers who 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.* 2013) but not with the current focus.

The choice experiment consisted of three alternatives; two labelled alternatives representing a cordon-based charging scheme and a distance-based charging scheme, randomly assigned to road pricing schemes 1 and 2 and the status quo. An illustrative choice screen, together with the boundaries of the proposed cordon-based charge area, is presented in Figure 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

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

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.

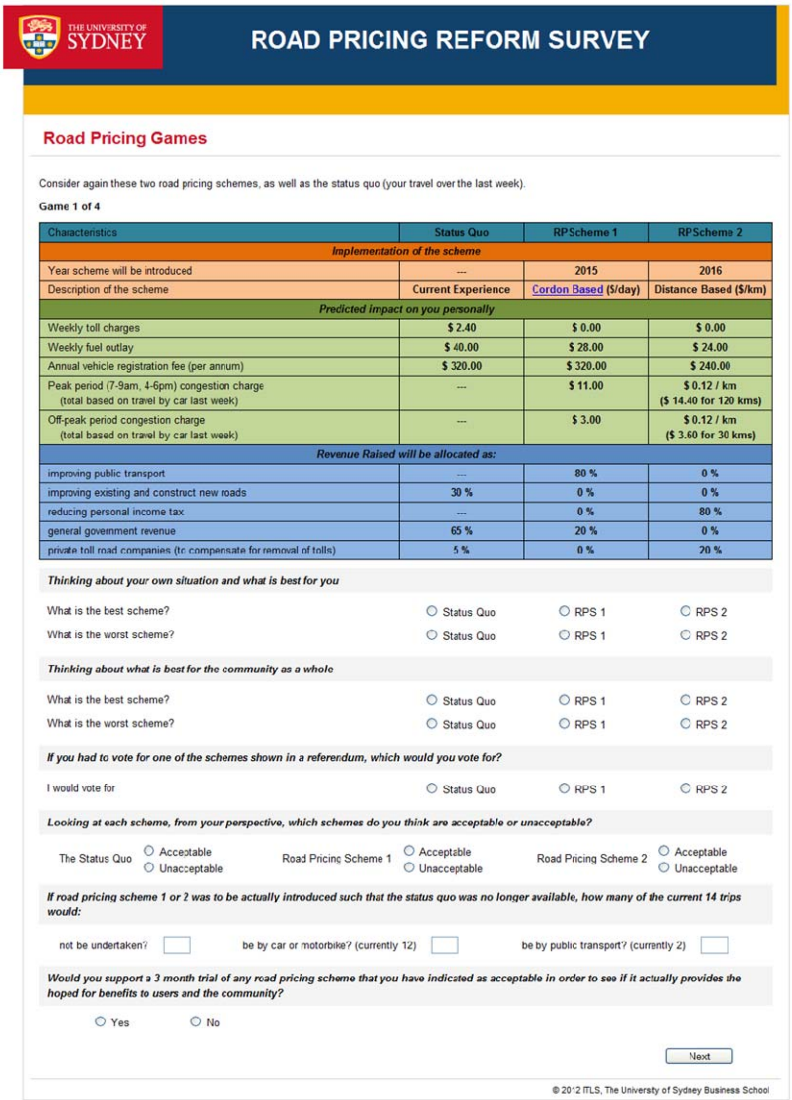
The allocation of revenues raised were fixed for the status quo alternative, but varied in the cordon-based and distance-based charging schemes over choice tasks. The allocation of revenue was varied from zero per cent to 100 per cent for a given revenue stream category. Within a charging scheme, the 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 per kilometre 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 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 four 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 of four choice tasks each. The blocking was accomplished by using an algorithm designed to minimise the maximum absolute correlation between the design attributes and the blocking column.

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%



THE UNIVERSITY OF SYDNEY

ROAD PRICING REFORM SURVEY

Road Pricing Games

Consider again these two road pricing schemes, as well as the status quo (your travel over the last week).

Game 1 of 4

Characteristics	Status Quo	RPScheme 1	RPScheme 2
Implementation of the scheme			
Year scheme will be introduced	---	2015	2016
Description of the scheme	Current Experience	Cordon Based ($/day)	Distance Based ($/km)
Predicted impact on you personally			
Weekly toll charges	$ 2.40	$ 0.00	$ 0.00
Weekly fuel outlay	$ 40.00	$ 28.00	$ 24.00
Annual vehicle registration fee (per annum)	$ 320.00	$ 320.00	$ 240.00
Peak period (7-9am, 4-6pm) congestion charge (total based on travel by car last week)	---	$ 11.00	$ 0.12 / km ($ 14.40 for 120 kms)
Off-peak period congestion charge (total based on travel by car last week)	---	$ 3.00	$ 0.12 / km ($ 3.60 for 30 kms)
Revenue Raised will be allocated as:			
improving public transport	---	80 %	0 %
improving existing and construct new roads	30 %	0 %	0 %
reducing personal income tax	---	0 %	80 %
general government revenue	65 %	20 %	0 %
private toll road companies (to compensate for removal of tolls)	5 %	0 %	20 %

Thinking about your own situation and what is best for you

What is the best scheme? ○ Status Quo ○ RPS 1 ○ RPS 2

What is the worst scheme? ○ Status Quo ○ RPS 1 ○ RPS 2

Thinking about what is best for the community as a whole

What is the best scheme? ○ Status Quo ○ RPS 1 ○ RPS 2

What is the worst scheme? ○ Status Quo ○ RPS 1 ○ RPS 2

If you had to vote for one of the schemes shown in a referendum, which would you vote for?

I would vote for ○ Status Quo ○ RPS 1 ○ RPS 2

Looking at each scheme, from your perspective, which schemes do you think are acceptable or unacceptable?

The Status Quo ○ Acceptable ○ Unacceptable Road Pricing Scheme 1 ○ Acceptable ○ Unacceptable Road Pricing Scheme 2 ○ Acceptable ○ Unacceptable

If road pricing scheme 1 or 2 was to be actually introduced such that the status quo was no longer available, how many of the current 14 trips would:

not be undertaken? [ ] be by car or motorbike? (currently 12) [ ] be by public transport? (currently 2) [ ]

Would you support a 3 month trial of any road pricing scheme that you have indicated as acceptable in order to see if it actually provides the hoped for benefits to users and the community?

○ Yes ○ No

Next

© 2012 ITLS, The University of Sydney Business School

Figure 1 The Location of the Cordon-Charge Area

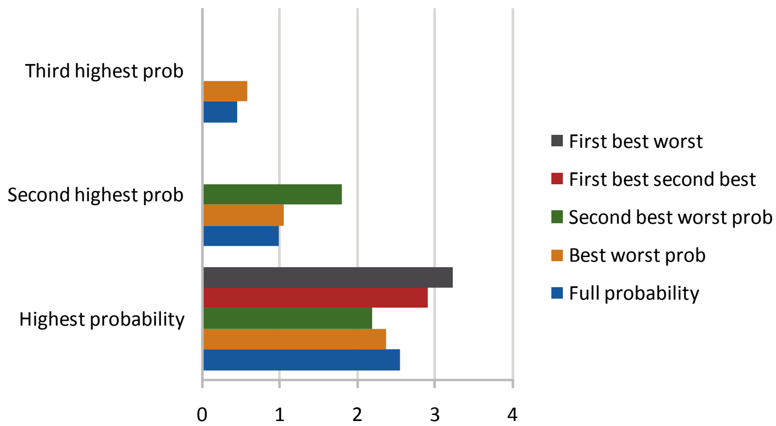
5. The Evidence

Mixed logit models were estimated in which the parameters associated with the four road pricing attributes were defined as random parameters with constrained triangular distributions[6]. The overall goodness of fit as the log-likelihood at convergence (as well as pseudo R^2 and Akaike information criteria (AIC)) for each of the four models is as follows: -648.77 (0.254, 1.667) for the full model (3 alternatives), -1353.63 (0.525, 1.712) for the 'exploded' best-worst model (3 and 2 alternatives), -432.76 (0.496, 1.127) for the 1st and 2nd best alternatives, and -338.70 (0.606, 0.892) for the 1st best and worst alternatives. The evidence suggests that the simple 2-alternative choice form of 1st best and worst is much preferred over all of the other processing forms.

Turning to the incidence of compliance with ISDA, in our application with four choice sets per respondent, ISDA is satisfied fully when we observe the chosen alternative having the highest probability in all four choice sets. The evidence is summarised in Table 2. It suggests for the full sample of 200 car users, that the specification that satisfies ISDA best (but not 100% across the sample) is the binary best-worst form (average of 3.24 out of 4). The full rank model (2.55/4) is an improvement over the exploded best-worst model (2.36/4); however as might be expected, the 1st best-2nd best form does not satisfy ISDA to the same extent (2.90/4) as the 1st binary best-worst, which is intuitively plausible because it narrows the difference in the choice probabilities, with an expectation that there will be more outcomes in which the best alternative from the rank is not the one with the greatest 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 ISDA than any other form assessed.

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

	Full probability	Best worst prob	Second best worst prob	First best second best	First best worst
Highest probability	2.56	2.37	2.20	2.90	3.24
Second highest prob	.99	1.05	1.80	-	-
Third highest prob	.46	.59	-	-	-



The findings in Table 2 can also be represented in terms of the average predicted probability of the 1st best alternative, summarised in Table 3. As expected, given the evidence in Table 2, the mean choice probability associated with the 1st best (rank order

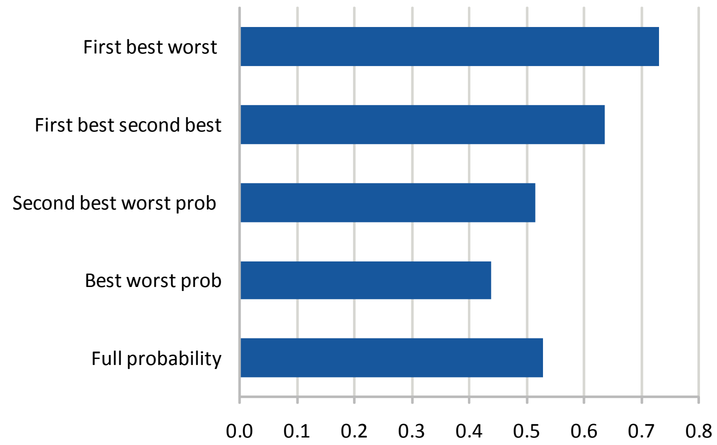
[6] The full models are available on request.

=1) varies from a high of 0.730 for the 1st best-worst regime, to a low of 0.437 for the traditional (exploded) best-worst regime. The range, close to 0.30 probability units, is significant and raises concerns about the behavioural validity of the traditional best-worst choice model form (noting however that the evidence is based on only one data set). The second best regime is the 1st best vs. 2nd best, which has diluted part of the probability gains associated with the 1st best in contrast to the worst alternative, as expected, since the 1st and 2nd best are clearly closer in preference space (i.e., utility terms). What is of special interest is the evidence suggesting that the full ranked model of three alternatives performs much better than the traditional best-worst regime.

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 both compliance with ISDA and the consequent significant improvement in the ability of the choice model to predict a respondent's 1st best alternative. Whether there are some exogenous influences at work that might increase our understanding of the differences in the evidence on incidence of compliance with ISDA is investigated in the following section.

Table 3 Average predicted probability of the best alternatives by model specification

	Full probability	Best worst prob	Second best worst prob	First best second best	First best worst
Average choice probability	0.52689	0.43714	0.51374	0.63642	0.73007



Before investigating sources of influence on incidence compliance under ISDA, a summary of an informative behavioural output, namely mean estimates of direct and cross elasticities for the road pricing reform pricing attributes associated with random parameters, is presented in Table 4. Even though the estimates are very small (highly inelastic), the important point to note is that the mean estimates are very different between the four choice processing regimes. The highest mean estimates for direct elasticities are for the full rank and the preferred choice regime of 1st best vs. worst, although the peak period cordon-based charge has similar direct elasticities to the 1st best vs. worst choice regime to the 1st best-2nd best regime. 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 choice making regimes. We would argue that compliance with ISDA might be an appealing guide in selecting the preferred choice making model, and hence the associated elasticity estimates.

Table 4 Direct and cross elasticities under alternative choice set conditions

	Full Rank	Exploded Best-Worst	1st Best and 2nd Best	1st Best and Worst
Number of alternatives	3	3 and 2	2	2
Peak period cordon charge				
Direct elasticities:				
SC alt 1	-0.0578	-0.0163	-0.0333	-0.0221
SC alt 2	-0.0464	-0.0133	-0.0252	-0.0199
Cross elasticities:				
SQ, Alt1	0.0311	0.0069	0.0169	0.0155
SQ, Alt2	0.0300	0.0059	0.0222	0.0113
Alt1, Alt2	0.0270	0.0057	0.0079	0.0176
Alt2, Alt 1	0.0287	0.0076	0.0097	0.0142
Off-peak period cordon				
Direct elasticities:				
SC alt 1	-0.0187	-0.0064	-0.0019	-0.0122
SC alt 2	-0.0177	-0.0058	-0.0016	-0.0138
Cross elasticities:				
SQ, Alt1	0.0126	0.0035	0.0009	0.0127
SQ, Alt2	0.0123	0.0035	0.0014	0.0057
Alt1, Alt2	0.0116	0.0027	0.0007	0.0154
Alt2, Alt 1	0.0115	0.0029	0.0010	0.0093
Peak period distance charge				
Direct elasticities:				
SC alt 1	-0.226	-0.0214	-0.0318	-0.2006
SC alt 2	-0.219	-0.0218	-0.0358	-0.1801
Cross elasticities:				
SQ, Alt1	0.0504	0.0129	0.0134	0.0418
SQ, Alt2	0.0464	0.0136	0.0090	0.0387
Alt1, Alt2	0.0449	0.0091	0.0149	0.0173
Alt2, Alt 1	0.0486	0.0083	0.0118	0.0292
Off-peak period distance				
Direct elasticities:				
SC alt 1	-0.1928	-0.0070	-0.0344	-0.1589
SC alt 2	-0.1677	-0.0059	-0.0430	-0.1148
Cross elasticities:				
SQ, Alt1	0.0373	0.0043	0.0189	0.0211
SQ, Alt2	0.0310	0.0040	0.0106	0.0233
Alt1, Alt2	0.0305	0.0026	0.0156	0.0116
Alt2, Alt 1	0.0366	0.0026	0.0122	0.0215

6. Systematic Factors linked to ISDA 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 ISDA compliance. To identify candidate influences, we ran a series of ordered logit models in which the dependent variable was the discrete compliance rate under ISDA for each respondent. The rate took the values 0 to 4 (see the frequency distribution given in Table 5). The results are summarised in Table 6, with partial (or marginal effects) available on request.

Table 5 Frequency Distribution of Compliance with ISDA

	Full Rank	BWexpl	B1B2	B1W
0.0	1.0%	4.5%	4.0%	2.0%
1.0	12.0%	21.0%	7.0%	4.5%
2.0	31.0%	23.0%	18.0%	12.5%
3.0	35.0%	33.5%	37.0%	29.5%
4.0	21.0%	18.0%	34.0%	51.5%
Total	100.0%	100.0%	100.0%	100.0%

A number of variables were identified as statistically significant influences on the incidence of compliance with ISDA. A particularly interesting result relates to the awareness of the road pricing debate (AWARE). The positive parameter associated with AWARE suggests that the incidence of compliance with ISDA increases as the perceived awareness response increases, highlighting the role that increased knowledge of the context of the theme of the choice experiment plays in ensuring that the stated preference ordering, translated into the choice probability (or preference) ordering, complies with ISDA.

Table 6 Sources of Explanation of Incidence of Compliance with ISDA (Parameter estimates with t-values in brackets)

	Full Rank	**Exploded Best-Worst**	**1st Best and 2nd Best**	**1st Best and Worst**
Number of alternatives	3	3 and 2	2	2
Sources of Influence:				
Constant	2.646 (36.4)	1.692 (16.6)	3.152 (28.2)	3.764 (41.2)
Age (years)		0.0066 (3.41)	-0.0037 (-1.96)	
Personal income ($000s)				-0.0031 (-4.82)
Gender (male=1)	-0.368 (5.76)	-0.252 (-4.14)	-0.2287 (-3.76)	-0.522 (-7.41)
Kms of travel per day		-0.0036 (-2.63)		0.0056 (3.35)
Annual registration fee ($)	0.0002 (2.25)	0.0036 (5.27)	0.0002 (2.45)	-0.0003 (-4.43)
Weekly fuel outlay ($)	0.0111 (9.82)	0.0063 (5.93)	0.0076 (7.57)	0.0132 (7.51)
Cordon-based peak period charge ($/day)	0.0105 (2.47)			0.0175 (3.87)
Distance-based charge in 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)
Awareness of road pricing debate (0-100)	0.3573 (3.25)	0.4885 (4.42)		0.6281 (5.80)
Threshold parameters:				
μ_1	1.498 (27.2)	1.295 (24.8)	1.0964 (17.2)	1.2464 (14.4)
μ_2	2.886 (53.4)	2.342 (42.0)	2.3033 (39.8)	2.5144 (33.7)
μ_3	4.506 (78.9)	3.971 (65.9)	3.889 (72.7)	3.994 (57.5)
Log-likelihood at convergence	-5686.79	-6022.34	-5379.20	-4554.43
Log-likelihood at zero	-5792.69	-6081.27	-5432.94	-4718.39
AIC	2.849	3.016	2.694	2.283

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 with ISDA increases under the exploded best-worst specification but decreases under the 1st best-2nd best choice processing rule. Personal income is only a statistically significant influence for the 1st best-worst processing rule, with the negative sign suggesting, *ceteris paribus*, that the incidence of compliance with ISDA decreases as income increases. Gender has a negative influence in all models, suggesting that compliance with ISDA 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 with ISDA 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 1st best-worst), 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 preference ordering complies more with ISDA. 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 with ISDA. Thus one might be tempted to conclude, given the available potential sources of systematic influence available in the data, that compliance of preference ordering with ISDA is aligned strongly with individual-specific idiosyncratic effects which are not revealed.

7. 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, focussing on the role that each alternative plays in establishing the preference ordering (and the *ex post* estimation associated choice probability) for an alternative.

Fundamentally, we question whether there is some redundancy in the offered set of alternatives[7] which may 'get in the way' of improving the capability of a model to predict the alternative that is stated as the preferred (i.e., 1st best) alternative. In this paper, theoretical (or behavioural) guidance is offered through ISDA in establishing a test of the relevance of each alternative, given the subjective rank ordered by a sample of respondents of each alternative in a predefined choice experiment and a predefined information processing rule.

[7] 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.

Under four proposed choice making rules, we have identified the extent to which the ISDA axiom is satisfied. While all processing strategies violate Quiggin's ISDA condition to some extent, the 1st 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 full rank choice set, the 1st best vs. 2nd 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, in future choice studies, the most appropriate processing rule in choice making. We encourage further testing using other data sets, especially where there are more than three alternatives offered in a choice experiment.

References

Collins, A. T., and Rose, J. M. (2011) Estimation of stochastic scale with best-worst data. Manuscript, Institute of Transport and Logistics Studies, University of Sydney.

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

Flynn, T. N., Louviere, J.J., Peters, T.J. and Coast, J. (2007) Best-Worst scaling: what it can do for health care research and how to do it, *Journal of Health Economics*, 26, 171–189.

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

Hensher, D.A. (2006) How do respondents process stated choice experiments? – Attribute consideration under varying information load, *Journal of Applied Econometrics*, 21 (6), 861-878.

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

Hensher, D.A., Rose, J.M. and Collins, A. (2013) Understanding buy in for risky prospects: incorporating degree of belief into the *ex ante* assessment of support for alternative road pricing schemes, *Journal of Transport Economics and Policy*, 47 (3), 453-73.

Louviere, J.J., Lings, I, Islam, T., Gudergan, S. and Flynn, T. (2013) An introduction to the application of (case 1) best–worst scaling in marketing research, *International Journal of Marketing Research*, 30, 292-303.

Marley, A. A. J., and Pihlens, D. (2012) Models of best-worst choice and ranking among multi-attribute options (profiles), *Journal of Mathematical Psychology*, 56, 24-34.

Rose JM (2014) 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., and Collins, A. T. (2008) Designing efficient stated choice experiments in the presence of reference alternatives, *Transportation Research Part B*, 42(4), 395-406.

Quiggin, J. (1982) A theory of anticipated utility, *Journal of Economic Behavior and Organization*, 3(4), 323-343.

Quiggin, J. (1995) Regret theory with general choice sets, *Risk and Uncertainty* 8(2), 153–65.