MPRA

Munich Personal RePEc Archive

A comparison of responses to single and repeated discrete choice questions

Ben J. McNair and Jeff Bennett and David A. Hensher

Crawford School of Economics and Government, The Australian National University

May 2010

THE AUSTRALIAN NATIONAL UNIVERSITY

Environmental Management & Development

A comparison of responses to single and repeated discrete choice questions

**#14
2010**

Ben J. McNair, Jeff Bennett and David A. Hensher

A comparison of responses to single and repeated discrete choice questions

Ben J. McNair[a,*], Jeff Bennett[a] and David A. Hensher[b]

[a] Crawford School of Economics and Government, The Australian National University, Canberra ACT 0200, Australia.

[b] Institute of Transport and Logistics Studies, Faculty of Economics and Business, The University of Sydney, NSW 2006, Australia.

* Corresponding author. Email address: ben.mcnair@anu.edu.au

Abstract   According to neoclassical economic theory, a stated preference elicitation format comprising a single binary choice between the status quo and one alternative is incentive compatible under certain conditions. Formats typically used in choice experiments comprising a sequence of discrete choice questions do not hold this property. In this paper, the effect on stated preferences of expanding the number of binary choice tasks per respondent from one to four is tested using a split sample treatment in an attribute-based survey relating to the undergrounding of overhead electricity and telecommunications wires. We find evidence to suggest that presenting multiple choice tasks per respondent decreases estimates of expected willingness to pay. Preferences

stated in the first of a sequence of choice tasks are not significantly different from those stated in the incentive compatible single binary choice task, but, in subsequent choice tasks, responses are influenced by cost levels observed in past questions. Three behavioural explanations can be advanced – weak strategic misrepresentation, reference point revision and cost-driven value learning. The evidence is contrary to the standard assumption of truthful response with stable preferences.

1. Introduction

Accurate estimation of social values for public goods and other non-market goods is crucial to ensuring improved welfare outcomes for public policy. Social value estimates are central to evaluations of the costs and benefits of public good provision, or, in cases where public or natural monopoly goods are provided by private firms, to setting performance-adjusted prices to internalise the social net benefit optimisation within a firm's profit-maximisation problem.[1] Discrete choice methods using stated choice experiment data, which were originally developed in the transport (Hensher and Truong, 1985) and marketing (Louviere and Woodworth, 1983) contexts, have become a popular approach to estimating social values for multiple-attribute non-market goods such as environmental goods (Bennett and Blamey, 2001) and monopoly service quality (Beenstock et al., 1998; Carlsson and Martinsson, 2008). The method typically involves presenting respondents with a sequence of choice tasks, where respondents indicate their preference between two or more attribute-based alternatives in each choice task. Recent literature has paid attention to the conclusion from mechanism design theory (Green and Laffont, 1979; Hurwicz, 1972; Mirrlees, 1971) that utility-maximising consumers may find it optimal to misrepresent their preferences in this survey format (Bateman et al., 2008b; Carson and Groves, 2007).

---

[1] The optimal social welfare outcome can be achieved by a quality-adjusted price cap with yardstick competition over both cost and quality (which is the culmination of theory developed by Spence (1975), Loeb and Magat (1979), Baron and Myerson (1982) and Shleifer (1985) among others).

Analysts seek to implement incentive compatible survey mechanisms in which truthful response is the utility-maximising response for all respondents regardless of their beliefs about others' responses (or, equivalently, truthful response is the dominant strategy for all respondents). It has long been recognised that a survey mechanism with a plurality social choice function can be incentive compatible if its elicitation format is a single binary (*SB*) choice between the status quo and an alternative (Farquharson, 1969).[2] A format comprising a sequence of binary choices between the status quo and various alternatives can only be incentive compatible where the social choice function is based on the random selection of a single choice task from each sequence (Carson and Groves, 2007). While this social choice function may be possible in a laboratory environment (Boyle et al., 2004), it lacks credibility in field surveys not least because respondents are unlikely to believe that the agency would discard the majority of the data that they expended resources collecting. As a result, the *SB* choice elicitation format is the only format that can be incentive compatible in field surveys.

The *SB* format has successfully been employed in the dichotomous choice contingent valuation (CV) context (where the cost of the alternative varies across respondents, but the good is fixed), particularly following its recommendation by the NOAA blue ribbon panel in 1993 (Arrow et al., 1993). However, there are numerous difficulties associated with employing the *SB* format in the attribute-based choice experiment context, where the

---

[2] A necessary condition is that the agency can credibly claim to be able to force any of the alternatives on any given respondent. However, it is not necessary for the *SB* choice survey to be binding (Carson et al., 1997) or a full public vote (Green and Laffont, 1978).

good in the alternative varies across respondents. Estimates of willingness-to-pay from *SB* data are less statistically significant than those from repeated choice data because of the absence of opportunities for institutional learning (Braga and Starmer, 2005) as well as the much lower number of choice observations. Some evidence suggests that estimating individual-specific taste intensities or the heterogeneity in taste intensities across a population using *SB* data is problematic (Rose et al., 2009). For these reasons, it is preferred, and in some cases necessary, to present multiple choice tasks per respondent when eliciting preferences for multiple attributes. Given that this format will continue to play an important role in value estimation, it is important to understand the extent to which the potential for strategic response identified in theory is realised in practice as a divergence between stated preferences and true underlying preferences.

This paper uses a split sample treatment of elicitation format in a web-based survey relating to the undergrounding of overhead electricity and telecommunications wires in the Australian Capital Territory to assess the effect on stated preferences of presenting multiple choice tasks per respondent as opposed to a single incentive compatible choice task. The elicitation formats employed in the survey include a *SB* choice task and a sequence of four binary choice tasks. In the latter repeated binary (*RB*) format, similar goods are offered at quite different cost levels over the course of a sequence. The objectives of this paper are to use the data from these two elicitation formats to:

1. test whether stated preferences are affected by presenting four as opposed to one attribute-based choice task per respondent;

2. test whether stated preferences in the first choice task presented are affected by advance knowledge that four as opposed to one choice tasks will be presented; and,

3. identify which of the existing theories of respondent behaviour are consistent with any difference in preferences stated in the two elicitation formats.

The paper is organised as follows. Section 2 discusses the effects of elicitation format on stated preferences identified in the literature both in theory and empirically. Section 3 sets out the design of the survey mechanism used in this study and the econometric modelling approaches used to analyse the data. The results of the analysis are set out in Section 4 and Section 5 concludes.

2. Elicitation format and stated preferences

Standard assumptions in discrete choice experiments have been that:

1. all respondents truthfully answer the question being asked; and,

2. true preferences are stable over the course of a sequence of questions.

These assumptions imply that value estimates derived from the *SB* and *RB* formats should be identical. It has been recognised that responses may become more accurate over the course of a sequence as respondents become more familiar with the choice task format. This process of institutional learning (Braga and Starmer, 2005) is thought to have the effect of reducing the variance of the random error component (or, equivalently,

increasing scale).[3] Conversely, fatigue is thought to decrease scale once respondents proceed beyond a certain point in a sequence of questions (Bradley and Daly, 1994; Caussade et al., 2005; Holmes and Boyle, 2005). While these effects suggest that scale may differ between the *SB* and *RB* formats, they do not imply a difference in value estimates derived from the two formats. There are numerous possible behavioural explanations for why value estimates derived from the two formats may differ. Each involves a violation of one or both of the standard assumptions set out above. We focus on four behavioural theories that have been raised in the literature and, in particular, the predictions of those theories where similar goods are offered at quite different costs over the course of a sequence.

The first theory we consider is *strategic misrepresentation*, which involves a violation of the first standard assumption. It has long been recognised in neoclassical economic theory that consumers may conceal their true preferences if it enables them to obtain a public good at a lower cost (Samuelson, 1954). In our case, this strategic misrepresentation would be manifest in the *RB* format by the rejection of an alternative that is preferred to the status quo when a similar good was offered at a lower cost in a previous choice task. In doing so, respondents increase the likelihood that their most preferred option across the sequence of choice tasks is implemented. Following Bateman et al. (2008b), we differentiate between *strong strategic misrepresentation*, in which respondents always reject a good if it was offered at a lower cost in a previous choice task, and *weak strategic*

---

[3] In the multinomial logit model, the scale parameter, λ, is an inverse function of the variance of the unobserved effects, $\sigma^2 = \pi^2/6\lambda^2$.

misrepresentation, in which respondents weigh up the rejection against the perceived risk of the good not being provided at the lower cost. The prediction in both cases is that derived willingness-to-pay (WTP) will be lower in the *RB* format than the *SB* format.

A second theory – *cost-driven value learning* – involves a violation of the second standard assumption. Plott (1996) argued that initial underlying preferences are often poorly formed and preferences are discovered as a respondent progresses through a sequence of choice tasks. We focus on a case in which preferences are formed based on the cost levels presented over the course of a sequence. The response behaviour is similar to that of the 'good deal / bad deal' heuristic described by Bateman et al. (2008b). An alternative is more (less) likely to be chosen if its cost level is low (high) relative to the levels presented in previous choice tasks. Such revisions could arise where respondents take some weighted average of cost levels presented in the sequence to that point as a signal for the quality of the good. If the weighting is asymmetric in terms of high and low cost levels, then WTP derived from the *RB* and *SB* formats may differ. Several similar theories have been termed anchoring or starting-point bias (Boyle et al., 1985; Flachaire and Hollard, 2007; Herriges and Shogren, 1996; Ladenburg and Olsen, 2008). In its simplest form, starting-point bias does not imply a difference in derived WTP from the two response formats since preferences are established prior to answering the first question (but after *observing* the first question).

DeShazo (2002) presents a theory of *reference point revision* in which preferences may be well-formed, but respondents' value functions shift when a non-status-quo option is chosen. The shift occurs because the selection of a non-status-quo option is viewed as a transaction up to a probability and this causes a revision of the reference point around

which the asymmetric value function predicted by prospect theory is centred (Kahneman and Tversky, 1979). The outcome is a reduced likelihood of acceptance when a lower-cost option has been accepted in a previous choice task. Only questions subsequent to the first in the *RB* format are affected, implying that derived WTP will be lower in the *RB* format than the *SB* format.

A fourth behavioural theory is that respondents may find some questions implausible and answer as though cost or attributes of the good are at levels considered more realistic or more likely (Carson and Groves, 2007). This represents a violation of the first standard assumption in the sense that respondents are answering different questions from those being asked. Bateman et al. (2008b) describe a *cost averaging strategy* in which respondents believe that the true cost that would be charged lies somewhere in the middle of the range of costs presented in a sequence. Under such a strategy we would expect alternatives with cost levels from the low (high) end of the range observed by the respondent to be accepted less (more) frequently than in a truthful response. This is the opposite effect to that observed under *strategic misrepresentation* or *cost-driven value learning*.

Turning to empirical evidence, several studies have tested the effects of expanding the *SB* format in the fixed good (CV) context. Recognising the large sample sizes required for statistically significant estimation when using the *SB* choice format, some CV surveys have incorporated a follow-up question in the elicitation format. Numerous studies have found differences in WTP implied by the first and second questions in this double-bounded CV format (Cameron and Quiggin, 1994; Hanemann et al., 1991; McFadden and Leonard, 1995). Several interpretations of the difference have been given. While the most

common appears to be some form of anchoring, DeShazo (2002) and Bateman et al (2008a) present evidence to support the theories of *reference point revision* and *value learning*, respectively. Evidence from Carson et al. (Carson et al., 2009) suggests the difference is due to *strategic misrepresentation*. They show that responses to the first and second questions are equivalent in the presence of a social choice function in which the outcome of the second question cancels out and replaces the outcome of the first question.

The attribute-based choice experiment format has not been subject to the same degree of testing for violations of the standard assumptions as the CV format.[4] Some studies have compared stated preferences from fixed-good (CV) *SB* and attribute-based repeated binary formats. For example, Cameron et al. (2002) were unable to reject the hypothesis

---

[4] A number of studies have focussed on hypothetical bias by comparing results from hypothetical choice experiments with those from choice experiments with immediate and certain implementation (Alfnes and Steine, 2005; Carlsson and Martinsson, 2001; Hensher, 2009; Lusk and Schroeder, 2004). Carson and Groves (2007) distinguish between inconsequential hypothetical surveys (where a respondent believes there is 0 per cent chance of implementation) and consequential hypothetical surveys (where a respondent believes their responses will influence up to some non-zero probability the likelihood of an alternative being implemented by the agency). The same conditions for incentive compatibility apply to the survey mechanism regardless of whether it is a consequential hypothetical survey or a survey with immediate and certain implementation. If the survey is inconsequential, then neoclassical economic theory cannot be used to predict responses. Consistent with this theory, Carson et al. (2006) found a difference between responses to inconsequential hypothetical questions and questions involving 100 per cent probability of actual payment, but, importantly, found equivalence in responses to all questions involving a non-zero (20 per cent, 50 per cent, 80 per cent and 100 per cent) probability of actual payment.

of identical indirect utility-difference functions across these elicitation formats. Several studies have examined the implications of presenting multiple attribute-based choice tasks per respondent without employing an incentive compatible *SB* comparator. Bateman et al. (2008b) found evidence of *weak strategic misrepresentation* using a split-sample treatment of advance knowledge of attribute levels in a sequence of choice tasks. Day and Pinto (2010) found evidence that stated preferences are influenced by both cost and non-cost attribute levels presented in previous choice tasks in a sequence. They conclude that choice behaviour is driven by some form of value learning. These studies, and others, have used the first choice task in a sequence as an 'unbiased' comparator. While the first choice task in a sequence is not influenced by information conveyed by other choice tasks, it does not possess the 'take it or leave it' property necessary for incentive compatibility.

The contribution offered by this research is two-fold. First, we make spilt-sample comparisons between preferences stated in a sequence of choice tasks and those stated in an incentive compatible attribute-based *SB* choice task. We are aware of just one existing study (Racevskis and Lupi, 2008) and one concurrent study (Scheufele and Bennett, 2010a) employing this approach. As yet, there is no body of evidence to support or counter Racevskis and Lupi's (2008) finding of a significant difference between models fitted to data from the two elicitation formats. This paper represents a contribution towards addressing this research gap. By comparing preferences stated in an incentive compatible *SB* choice format with those stated in the first choice task of a sequence, we improve the understanding of whether the first question in a sequence can be used as an

unbiased comparator (thus saving the considerable expense of collecting a single choice observation per respondent).

Second, we model the relationship between cost sensitivity and the positioning of the cost level relative to levels presented in previous choice tasks. A number of authors, including Bateman et al. (2008b) and Carson and Groves (2007), have discussed the potential for such effects, but few studies have modelled them. Holmes and Boyle (2005) model the effects of cost levels in previous and future choice tasks using data from a paper-based survey. Day and Pinto (2010) use non-parametric methods on data from split-samples presented with sequences of choice tasks with increasing and decreasing cost levels. We take a different approach to Day and Pinto (2010) by using econometric models to control for ordering anomalies. Our model differs from Holmes and Boyle (2005) in that it is tailored to data from an internet-based survey that prevents backward navigation through the sequence of choice tasks. Rather than specifying cost differences with lead and lag choice tasks as drivers of behaviour, we use categories representing the positioning of the cost level relative to the full history of cost levels presented to the respondent. As far as possible, we relate the results to the behavioural theories outlined above.

3. Research design and method

The empirical testing was carried out on data from a survey of homeowners in the Australian Capital Territory (ACT) in 2009. The main objective of the survey was to establish homeowners' willingness to pay to have overhead electricity and telecommunications wires in their suburb replaced by new underground wires. This

process of undergrounding confers a number of benefits on households. Underground networks generally provide a more secure and reliable service. They are less prone to damage from fires, strong winds, ice storms, lightning and other severe weather events, which can lead to extended power outages and risks of electrocution. They lead to more aesthetically pleasing residential areas and they avoid costs associated with trimming trees away from overhead power lines. Efforts to conduct economic evaluations of potential underground projects have been hampered by a lack of knowledge about the value households place on these benefits. For example DCITA (1998), Infrasource (2007) and IPART (2002) simply categorise most of these benefits as unquantifiable. The estimates derived from this study therefore represent valuable information for policy-makers considering the economic merits of undergrounding where little or no information was previously available.

The survey employed a hybrid stated preference methodology, combining the attribute-based approach of choice experiments with the project-based dichotomous choice approach of contingent valuation. The elicitation formats employed in the survey included a single binary choice task (*SB*) and a sequence of four binary choice tasks (repeated binary, *RB*).[5] In each binary choice task, respondents were presented with a description of their current (overhead) service and one undergrounding alternative.[6] The

[5] A third elicitation format, a sequence of four choice tasks containing two alternatives to the status quo (RMN), was also included but data from this format is not analysed in this paper.

[6] All non status quo alternatives involved new underground infrastructure. This ensured that every alternative in the design was meaningful as a *SB* choice, while allowing the same set of alternatives to be used in all elicitation formats.

attributes used to describe the alternatives and the levels assigned to those attributes are presented in Table 1. All of the benefits of undergrounding other than supply reliability benefits are embodied in the alternative label, including the amenity and safety benefits that qualitative questions showed to be the major household benefits from undergrounding. Two blocks of four choice tasks were constructed in a format comprising the status quo alternative plus two undergrounding alternatives to maximise the Bayesian C-efficiency of the design (Scarpa and Rose, 2008) and minimise the correlation between attribute levels and block assignment.[7] The *RB* design was created by splitting these two blocks into four blocks of four *binary* choice tasks. The *SB* design is an extreme blocking of the *RB* design, with each respondent in the *SB* sample split receiving one of these 16 choice tasks.

The web-based questionnaire was refined based on in-depth interviews with 11 participants. Households were recruited by telephone and screening questions were used to ensure that participating households were owner-occupiers of stand-alone houses serviced by overhead wires. Email invitations were sent to the 2,485 households that agreed to participate. 1,744 respondents completed the online questionnaire; 1,163 in the *SB* sample split and 292 in the *RB* sample split. Table 2 shows there is no significant difference in the socio-demographic composition of the two groups.

---

[7] Bayesian priors were derived from pilot responses and from NERA and ACNielsen (2003). Default levels were assumed for supply reliability attributes in the status quo. The RMN design was used because it was expected that estimates of WTP for supply reliability, which were less statistically significant in the design than the alternative label and cost attribute, would rely heavily on data from that elicitation format.

Table 1: Attributes and levels

Attribute	Levels	
	Status quo (overhead) alternative	Undergrounding alternatives
Your one-off undergrounding contribution (AUD 2009)	0	1,000, 1,100, 2,000, 2,100, 2,800, 3,000, 3,900, 4,000, 6,000, 6,200, 8,000, 8,200, 11,800, 12,000, 15,900, 16,000
Power cuts without warning:		
Number of power cuts each five years	Set by respondent	Proportions of status quo level: 0.25, 0.5, 0.75, 1 [a,b]
Average duration of power cuts	Set by respondent	Proportions of status quo level: 0.33, 0.66, 1.33, 1.66 [a]
Power cuts with written notice (occurring in normal business hours):		
Number of power cuts each five years	Set by respondent	Proportions of status quo level: 0.2, 0.4, 0.6, 0.8 [a,b]
Average duration of power cuts	Set by respondent	Proportions of status quo level: 0.33, 0.66, 1.33, 1.66 [a]

[a] Rounded to the nearest integer; [b] Absolute levels (0, 1 and 2) were assigned where respondents chose very low status quo levels (1 or less).

The questionnaire described the research project as a partnership between the local

electricity network service provider, ActewAGL Distribution, and the two universities.

After establishing individual-specific reference levels for power supply reliability

attributes, the questionnaire advised respondents of the number of choice tasks that would

be presented, the number of alternatives that would be presented in each task and the

attributes that would be used to describe each alternative. The questionnaire outlined a

suburb-based majority rule social choice function (often referred to as a provision rule or

decision rule in the non-market valuation literature) that ensured incentive compatibility

in the *SB* response format. In the *RB* format, the equivalent social choice function was

that an undergrounding option would be *considered* for implementation in a suburb if it was preferred to the status quo by more than 50 per cent of respondents in that suburb.

Importantly, the survey instrument did not allow respondents to navigate back through the sequence of choice tasks. The survey was programmed to cycle through the various blocks, choice task orderings and elicitation format sample splits to ensure approximately equal representation across choice observations. Data were excluded from the models presented in this paper where respondents took less than five minutes to complete the *SB* survey or less than six minutes to complete the *RB* survey. It was judged that these responses were given without consideration (possibly randomly) solely as a means of qualifying for the prize draw participation incentive.

Table 2: Socio-demographic statistics by sample split

Variable	*SB* sample mean	*RB* sample mean	Test of difference in groups (p-value)[a]
Household size (persons)	2.94	3.06	0.402
Gender (% male)	51.9	53.1	0.703
Age:			
% under 40	18.8	19.7	0.733
% over 65	15.7	15.2	0.832
Highest level of education:			
% undergraduate degree	29.4	26.6	0.343
% postgraduate degree	31.2	34.1	0.333
Annual household income:			
% under $52,000	11.3	10.0	0.533
% over $104,000	39.2	39.0	0.944
% refused	18.2	17.6	0.816

[a] p-values calculated from χ^2-test of differences in proportions across groups, except p-value for household size, which is calculated from a two-sided t-test.

Two models are used to address the research objectives – the binary logit (BL) model and the random parameters logit (RPL) model. Both models are based on random utility theory (McFadden, 1980), which is built on the assumption that the utility, U, derived by a respondent from an alternative is a function of the attributes of the alternative, choice invariant characteristics (such as the characteristics of the respondent or the choice task) and a random element, ε. In any given choice task, respondents choose the alternative that yields the highest utility. The outcome is an index of the observed choice, y. The utility that respondent i derives from alternative j in choice task t is $U_{ijt} = \alpha_{ij} + \beta_i' \mathbf{x}_{itj} + \varepsilon_{itj}$ where $\mathbf{x}_{itj}$ is a vector of observed variables, α_{ij} is an alternative specific constant taking the value zero in the status quo alternative and β_i is a vector of coefficients to be estimated. The assumption that ε is independently and identically distributed according to the extreme value type I function gives the logit model form. In the models herein, all choice tasks comprise two alternatives and all observed variables are defined in such a way that $\mathbf{x}_{it} = 0$ in the status quo alternative ($j=1$). This allows the RPL model choice probability function for respondent i in choice task t to be written:

$$\pi_{it} = \Pr\left(y_{it} = 1 \mid x_{it}, z_{it}, v_i\right) = \frac{\exp(\alpha + \beta_{it}' x_{it})}{1 + \exp(\alpha + \beta_{it}' x_{it})} \tag{1}$$

where

$\beta_{it} \qquad = \beta + \delta \mathbf{z}_{it} + \Gamma \Omega_{it} \mathbf{v}_i$

$\mathbf{z}_{it} \qquad$ = a vector of choice task characteristics (such as its position order in a sequence)

$\Omega_{it} \qquad$ = a diagonal matrix of choice task specific variance terms; $\omega_{it} = \exp(\omega_t' \mathbf{z}_{it})$

$\mathbf{v}_i$ = a vector of random variables with zero mean and known variances

Γ = the lower triangular Cholesky matrix containing rows of zeros where

parameters are non-random

The models presented herein are based on the assumption that all attributes are strictly

processed as full compensatory. The parameters to be estimated are α, the coefficient

vectors β and δ and the non-zero elements of Γ and Ω_{it}. The BL model is a special case in

which $\Gamma = 0$ (and $\Omega_{it} = 0$).

To test whether stated preferences are affected by presenting four as opposed to one

attribute-based choice task per respondent, we compare expected WTP for the mean

undergrounding scenario from a BL model estimated on data from the *SB* format with the

equivalent estimate from a BL model estimated on data from the *RB* format. Expected

WTP is a Hicksian compensating measure of welfare change calculated analytically as

the area under the choice probability function truncated at the maximum cost level

($16,000) with all non-cost variables set at their population means:[8]

$$E(WTP) = \int_0^{x_{\max}^{\cos t}} \pi_{it} dx^{\cos t} = \int_0^{x_{\max}^{\cos t}} \frac{\exp(\overline{W}_{it})}{1 + \exp(\overline{W}_{it})} dx^{\cos t} \tag{2}$$

where

$$\overline{W}_{it} = \alpha + \beta_{it}' \overline{x}_{it} + \beta_{it}^{\cos t} x_{it}^{\cos t}$$

---

[8] The analytical calculation of the integral in Small and Rosen's (1981) equation 5.5 is required as the log

transformation on cost in our models prevents the use of the well-known explicit evaluation (Small and

Rosen's equation 5.9).

$\bar{x}_{it}$ = a vector of the means of non-cost observed variables

We restrict our analysis of the *SB* choice data to the BL model because models estimating heterogeneity in taste (RPL models), scale (scaled multinomial logit model) or both (generalised mixed logit model) across individuals are problematic when estimated on data with a single choice observation per respondent.[9] The BL model is estimated on the *RB* data to enable direct comparison.

To test whether stated preferences in the first of a sequence of choice tasks are affected by advance knowledge that multiple choice tasks will be presented, we incorporate in the BL model on the *RB* data interactions between the cost variable and effects coded variables for the order in which choice tasks were presented (with variables indicating the first three order positions, *q1*, *q2* and *q3*, taking the value -1 in the fourth order position). We again compare WTP for the mean undergrounding scenario from BL models on the two response formats, but here we evaluate WTP at the first question in the sequence in the model on the *RB* data.

To identify behavioural explanations for any differences in stated preferences between the two response formats, we focus on the influence of cost levels presented in previous choice tasks. We incorporate, in both the BL and RPL models on the *RB* data, variables

---

[9] Although mixed logit models may be able to disentangle the Gumbel error distribution and the random parameter distributions when estimated on repeated choice data (Fosgerau and Nielsen, 2006), further work is required to establish whether this is true of models estimated on single binary choice data. Rose et al (2009) found random parameter estimates statistically insignificant where data were a single choice observation per respondent in their study of the impact of the number of choice tasks per respondent.

accounting for the four possible 'relative positions' for the cost level presented in a

choice task.[10] In any given choice task, the cost level presented must be either:

a)   both the minimum and the maximum level presented in the sequence to that point

        (*m11*);

b)   the minimum, but not the maximum level presented in the sequence to that point

        (*m10*);

c)   the maximum, but not the minimum level presented in the sequence to that point

        (*m01*); or

d)   neither the minimum nor the maximum level presented in the sequence to that

        point (*m00*).

The full sample of *RB* data contains 289, 299, 297 and 271 choice observations for the

m11, *m10*, *m01* and *m00* relative cost positions, respectively. The *m11*, *m10* and *m01*

indicator variables are effects coded, taking the value -1 when relative cost position is

m00. These are interacted with the cost variable to allow estimation of the relationship

between cost sensitivity and the positioning of the cost level relative to the levels

presented to the respondent in previous choice tasks. For example, if the parameter

estimate for the *m00* interaction is significantly higher than that for the *m01* interaction,

this indicates that cost sensitivity is lower (and WTP is higher) when the cost level is

---

[10] The order variables are omitted from this model since the *q1* and *m11* variables are essentially identical

(since the cost level in first choice task presented is always both the minimum and maximum presented to

that point and this is not possible at any other point in the sequence).

within the range of levels presented in previous choice tasks relative to when it is the

highest level presented in the sequence to that point (with all other variables, including

cost, held constant). This would imply, for example, that an alternative with a cost level

of $4,000 is more likely to be chosen if previously presented cost levels were $2,000 and

$6,000 than if they were $2,000 and $1,000.

The relationships between the relative cost position interaction coefficients implied by the

various behavioural theories are presented in Table 3. *Strong strategic misrepresentation*

predicts an increase in cost sensitivity when the cost level is not the minimum presented

in the sequence to that point. Alternatives are rejected whether cost is the highest level

observed by the respondent or within the range observed. Under *weak strategic*

misrepresentation the relationship between these two relative positions is uncertain since

likelihood of acceptance depends on the maximum level of cost *accepted* in the sequence

to that point. The parameter relationships implied by *reference point revision* are the

same as those implied by *weak strategic misrepresentation*, with likelihood of acceptance

decreasing where a lower-cost alternative has been presented in an earlier choice task.

Cost-driven value learning implies that cost sensitivity is increased when cost is the

highest level observed by the respondent and decreased when cost is the lowest level

observed by the respondent relative to the first question or to questions in which cost lies

within the range observed by the respondent. *Cost averaging* implies the opposite set of

effects. The standard assumptions imply that relative cost position has no influence.

Using these implied relationships we examine whether evidence from our models is

consistent with or counter to the various theories outlined in Section 2. Our experiment

was not designed to test each behavioural hypothesis, so the identification of a single

overriding theory is unlikely. However, some discrimination is possible and this allows us to narrow the range of potential explanations for the response behaviour witnessed in the *RB* sample.

Table 3: Parameter relationships implied by behavioural theories

Behavioural theory	Implied parameter relationships
Strong strategic misrepresentation	$\beta_{m11*cost} = \beta_{m10*cost} > \beta_{m01*cost} = \beta_{m00*cost}$
Weak strategic misrepresentation	$\beta_{m11*cost} = \beta_{m10*cost} > \beta_{m01*cost}$ and $\beta_{m11*cost} = \beta_{m10*cost} > \beta_{m00*cost}$
Cost-driven value learning	$\beta_{m10*cost} > \beta_{m11*cost} > \beta_{m01*cost}$ and $\beta_{m10*cost} > \beta_{m00*cost} > \beta_{m01*cost}$
Reference point revision	$\beta_{m11*cost} = \beta_{m10*cost} > \beta_{m01*cost}$ and $\beta_{m11*cost} = \beta_{m10*cost} > \beta_{m00*cost}$
Cost averaging	$\beta_{m10*cost} < \beta_{m11*cost} < \beta_{m01*cost}$ and $\beta_{m10*cost} < \beta_{m00*cost} < \beta_{m01*cost}$
Truthful response with stable preferences	$\beta_{m10*cost} = \beta_{m11*cost} = \beta_{m00*cost} = \beta_{m01*cost} = 0$

4. Results

The BL model results are summarised in Table 4. The basic BL models on the *SB* and *RB* formats (Models 1 and 2, respectively) include a constant, the natural log of the household contribution (cost) and supply reliability attributes. The log transformation of the cost variable is utilised because it results in a better model fit. The choice probability (or bid acceptance) curves derived from Models 1 and 2 with non-cost variables set at their population means are shown in Figure 1. Bid acceptance is significantly lower in the *RB* format relative to the incentive compatible *SB* format for all cost levels except those at the lower end of the range used in the design. Estimates of mean WTP for the mean

undergrounding scenario derived from these models are presented in Table 5 along with confidence intervals derived from a bootstrapping procedure with 1000 random draws (calculated by drawing from normal distributions for all relevant parameters with moments set at their means and standard errors). Point estimates for mean WTP are $6,908 and $5,369 in the BL models on the *SB* and *RB* data, respectively. A one-sided test based on random draws of 1000 paired differences has a p-value of 0.0775. At the 90 per cent confidence level we reject the null and accept the alternative hypothesis that WTP is lower in the *RB* format than in the incentive compatible *SB* format.

Figure 1: Bid acceptance curves derived from Models 1 and 2
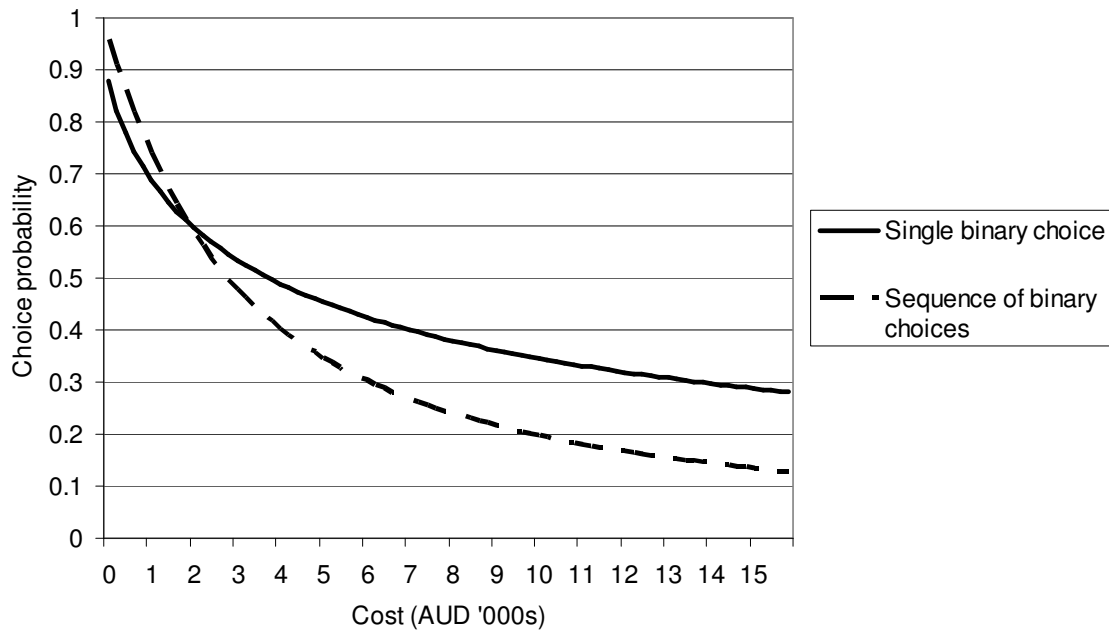
Table 4: Summary of results from binary logit models

Model	Model 1	Model 2	Model 3	Model 4	Model 5
Response format	*SB*	*RB*	*RB*	*RB*	*RB*
Parameter estimates:					
Constant	.68486***	1.00020***	1.02303***	.73865***	.74583***
	(.15256)	(.15343)	(.15466)	(.16645)	(.16589)
Log of household contribution	-.66964***	-1.14726***	-1.17542***	-.91672***	-.92002***
($'000s)	(.08050)	(.09017)	(.09177)	(.10940)	(.10922)
Change in number of unplanned	-.05588	-.04107	-.04281	-.07677*	-.07391*
outages per 5 years	(.05505)	(.04210)	(.04248)	(.04478)	(.04433)
Change in unplanned minutes off	-.00060	-.00010	-.00011	.00005	.00004
supply per 5 years	(.00055)	(.00046)	(.00046)	(.00048)	(.00047)
Change in number of planned	-.09450	-0.17447*	-.18259**	-.15965*	-.15752*
outages per 5 years	(.08311)	(.08903)	(.08992)	(.08976)	(.08932)
Change in planned minutes off	-.00009	-.00031	-.00029	-.00024	-.00023
supply per 5 years	(.00029)	(.00026)	(.00027)	(.00025)	(.00025)
Interactions with log of household contribution:					
Order: question 1 (q1=1)			.22681***	-.05439	
			(.06788)	(.17492)	
Order: question 2 (q2=1)			-.08210	-.04245	
			(.07511)	(.08108)	
Order: question 3 (q3=1)			-.12044	-.00607	
			(.07698)	(.10942)	
Minimum cost in sequence to that				.27262***	
point (minimum=1)				(.10032)	
Maximum cost in sequence to that				-.10694	
point (maximum=1)				(.06756)	
Relative cost position: m11					.11544
					(.07307)
Relative cost position: m10					.35776***
					(.11826)
Relative cost position: m01					-.36935***
					(.08038)
Model fit:					
Observations	1090	1112	1112	1112	1112
Log-likelihood	-711	-645	-639	-630	-630
Information criterion AIC	1433	1301	1296	1282	1279

*, ** and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels respectively; standard errors are in parentheses.

Table 5: Estimates of WTP for mean undergrounding scenario (AUD 2009)

|	Mean	95 per cent CI
Model 1 (single binary choice)	6,908	5,212 – 8,595
Model 2 (repeated binary choices)	5,369	4,058 – 6,819
Model 3 evaluated at:		
Question 1	6,586	4,843 – 8,389
Question 2	4,929	3,595 – 6,491
Question 3	4,786	3,483 – 6,257
Question 4	5,236	3,635 – 7,305
Model 5 evaluated at:		
Cost minimum and maximum presented	6,543	4,587 – 8,629
Cost minimum, but not maximum presented	8,116	5,632 – 10,583
Cost maximum, but not minimum presented	4,140	2,910 – 5,604
Cost neither minimum nor maximum presented	5,353	3,397 – 7,963

Turning to the second research objective, Model 3 incorporates interactions between cost and effects coded variables for the order in which choice tasks were presented to respondents in the *RB* format. This model provides separate estimates of cost sensitivity at each of the four order positions in the sequence, where cost sensitivity is defined as:

$$CostSens_t = -\partial U/\partial cost = -(\beta_{cost} + \beta_{q1*cost}.q1_t + \beta_{q2*cost}.q2_t + \beta_{q3*cost}.q3_t) \qquad (3)$$

The parameter estimate for the *q1* interaction is positive and significantly higher than the parameter estimates for the other question order interactions. This indicates that cost sensitivity is significantly lower in the first question relative to the later questions in the sequence with all other variables held constant. The modelled relationship between

question order and WTP (mean and 95 per cent confidence interval) is presented in

Figure 2. At $6,586, the estimate of mean WTP in the first choice task is similar to the

point estimate of mean WTP from the incentive compatible *SB* format of $6,908. We fail

to reject the null hypothesis of equivalence in a two-sided test based on 1000 randomly

drawn paired differences with a p-value of 0.8155. That is, we find no evidence to

suggest that advance knowledge that multiple choice tasks would be presented has an

effect on stated preferences in the first choice task. Therefore, the difference in WTP

estimates from the two response formats is driven by what is happening in the second,

third and fourth questions in the sequence. In particular, these results suggest cost

sensitivity increases (and derived WTP decreases) after the first question has been

answered.

Figure 2: WTP by question order (with 95 per cent confidence intervals)
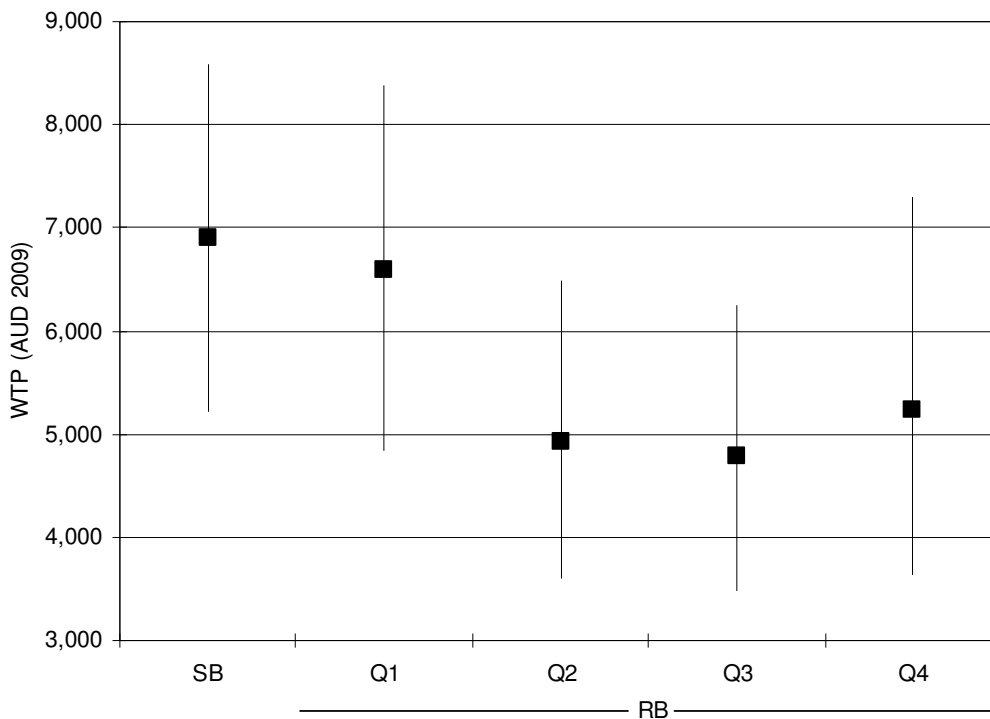
We turn now to the identification of behavioural explanations for this result. In Model 4, the question order variables are retained and *Min* and *Max* interactions are added where *Min* and *Max* are (1,-1) variables indicating whether the cost level is the minimum and maximum, respectively, presented to the respondent in the sequence to that point. The positive coefficient on the *Min* interaction indicates that cost sensitivity is lower when cost is the minimum presented in the sequence to that point with all other variables, including cost, held constant. The AIC value indicates an improvement in the explanatory power of the model. All of the effects coded variables for question order become insignificant indicating that there is no significant residual order effect once we account for the effect of cost levels presented in earlier choice tasks. This is a key result because it indicates that other possible interpretations for order effects, such as learning and fatigue, do not appear to be significantly affecting WTP or cost sensitivity in this case, though they may be affecting the error variance (and scale).

The effects of cost levels presented in previous choice tasks are examined in more detail using Model 5, which includes interactions between cost and effects coded variables for relative cost position. The model provides separate estimates of cost sensitivity at each of the four relative positions, where cost sensitivity is defined as:

$$CostSens_t = -\partial U/\partial cost = -(\beta_{cost} + \beta_{m11*cost}.m11_t + \beta_{m10*cost}.m10_t + \beta_{m01*cost}.m01_t) \quad (4)$$

The coefficient estimate on the *m10* interaction is highest, which indicates that cost sensitivity is lowest and WTP is highest when cost is the minimum, but not the maximum level presented in the sequence to that point (holding constant all other variables, including cost). The second highest is the coefficient estimate on the *m11* interaction,

followed by the implicit coefficient estimate on the *m00* interaction, which is equal to the negative of the sum of the three estimated coefficients. The coefficient estimate on the *m01* interaction is lowest, indicating that cost sensitivity is highest when cost is the maximum, but not the minimum level presented in the sequence to that point. Table 6 presents one-sided t-tests of the statistical significance of the differences. The null hypotheses are rejected and the alternative hypotheses of ordering consistent with the point estimates are accepted at the 90 per cent confidence level. At the 95 per cent confidence level, we fail to reject only one of the null hypotheses: that the coefficient on the *m10* interaction is not greater than the coefficient on the *m11* interaction.

Table 6: T-tests for differences in coefficients on relative cost position interactions

H_0:	$\beta_{m10*cost} - \beta_{m11*cost} \leq 0$	$\beta_{m10*cost} - \beta_{m00*cost} \leq 0$	$\beta_{m10*cost} - \beta_{m01*cost} \leq 0$	$\beta_{m11*cost} - \beta_{m00*cost} \leq 0$	$\beta_{m11*cost} - \beta_{m01*cost} \leq 0$	$\beta_{m00*cost} - \beta_{m01*cost} \leq 0$
p-value	0.0754	0.0029	0.0000	0.0251	0.0000	0.0129

At the 95 per cent confidence level, the relationships between the relative cost position interactions are consistent with the *weak strategic misrepresentation* and *reference point revision* theories, which both imply $\beta_{m11*cost}=\beta_{m10*cost}>\beta_{m01*cost}$ and $\beta_{m11*cost}=\beta_{m10*cost}>\beta_{m00*cost}$. Both of these theories explain the difference in WTP derived from the *RB* and *SB* formats. At the 90 per cent confidence level, the parameter relationships are consistent with the *cost-driven value learning* theory ($\beta_{m10*cost}>\beta_{m11*cost}>\beta_{m01*cost}$ and $\beta_{m10*cost}>\beta_{m00*cost}>\beta_{m01*cost}$). However, in order for *cost-driven value learning* to adequately explain the difference in WTP derived from the two elicitation formats, the behaviour would need to be asymmetric, with lower cost levels causing greater preference revisions than higher cost levels. The results are clearly

contrary to the *cost averaging* theory ($\beta_{m10*cost} < \beta_{m11*cost} < \beta_{m01*cost}$ and

$\beta_{m10*cost} < \beta_{m00*cost} < \beta_{m01*cost}$) and *truthful response with stable preferences*

($\beta_{m10*cost} = \beta_{m11*cost} = \beta_{m00*cost} = \beta_{m01*cost} = 0$).

Figure 3: WTP by relative cost position (with 95 per cent confidence intervals)
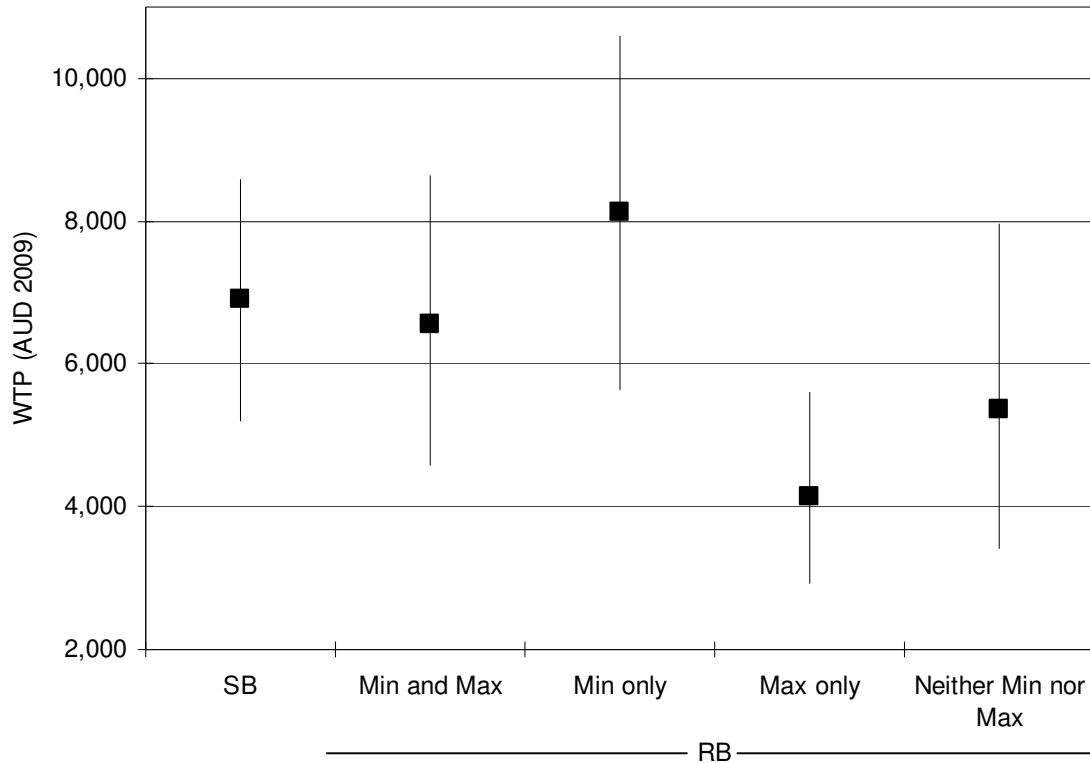
The modelled relationship between relative cost position and WTP is presented in Figure

3. A necessary and sufficient condition for the *m11* relative cost position is that the

question is the first in the sequence. Consistent with this, the WTP estimate evaluated at

the *m11* relative cost position of $6,543 is similar to the estimate from the incentive

compatible *SB* choice data of $6,908. We fail to reject the null hypothesis of equivalence

in a two-sided test based on 1000 randomly drawn paired differences with a p-value of

0.8024. The non-zero estimates of WTP when evaluated at the *m01* and *m00* relative

positions suggest that *strong strategic misrepresentation* was not widely employed.

In the RPL model (Model 6) on the *RB* choice data presented in Table 7 we use the

relative cost position variables to describe heterogeneity in the mean and

heteroscedasticity of a random cost parameter. The outcomes are estimated distributions

of cost sensitivity across the population evaluated at each of the relative cost positions

(see Figure 4),[11] where cost sensitivity is defined as:

$$CostSens_{it} = -\partial U/\partial cost$$

$$= -\beta_{it}^{cost}$$

$$= -[\beta^{cost} + \beta^{m11*cost}.m11_t + \beta^{m10*cost}.m10_t + \beta^{m01*cost}.m01_t + (\sigma^{cost} +$$

$$\sigma^{m11*cost}.m11_t + \sigma^{m10*cost}.m10_t + \sigma^{m01*cost}.m01_t).v_i], \quad v_i \sim N[0,1] \qquad (5)$$

The result of interest is the distribution of cost sensitivity when cost is the minimum and

not the maximum presented in the sequence to that point (*m10*). Mean cost sensitivity is

reduced relative to the first question (*m11*) and the variance in cost sensitivity is

significantly increased (one-sided t-tests have p-values 0.0355 and 0.0001, respectively).

This indicates there is significant heterogeneity in response behaviour when cost is lower

than the levels previously observed. The combination of the shift and the change in shape

of the distribution indicates that most respondents answered *m10* questions differently

than if the question had been presented as the first in the sequence, but to differing

---

[11] The distributions imply that some respondents exhibit a behaviourally implausible positive cost coefficient, which is not uncommon when using unconstrained random parameters, particularly when adding variables to describe heterogeneity around the mean (Hensher and Greene, 2009; Hensher et al., 2005).

degrees. This evidence supports a *cost-driven value learning* theory in which the magnitude of upward preference revision varies across respondents. Alternatively, it supports a mixture of behaviours across respondents, dominated by *cost-driven value learning* and one or both of *weak strategic misrepresentation* and *reference point revision*.

Table 7: Summary of results from random parameter logit model

Model	Model 5	
Response format	*RB*	
Random parameter: mean		
Log of [1 + household contribution ($'000s)]	-3.86793***	(0.84452)
Non-random parameters:		
Change in number of unplanned outages per 5 years	0.10437	(0.10608)
Change in unplanned minutes off supply per 5 years	-0.00122	(0.00132)
Change in number of planned outages per 5 years	-0.62662**	(0.26552)
Change in planned minutes off supply per 5 years	-0.00057	(0.00073)
Alternative specific constant (undergrounding = 1)	5.36297***	(0.97546)
Random parameter: heterogeneity in mean		
Relative cost position: m11	-0.43187	(0.44391)
Relative cost position: m10	2.31286**	(1.13387)
Relative cost position: m01	-1.25095**	(0.59936)
Random parameter: standard deviation		
Log of household contribution ($'000s)	3.64465***	(0.53801)
Random parameter: heteroscedasticity		
Relative cost position: m11	-0.18012	(0.15211)
Relative cost position: m10	1.10898***	(0.26896)
Relative cost position: m01	-0.42847**	(0.20356)
Model fit:		
Observations	1112	
Log-likelihood	461	
Information criterion AIC	947	

*, ** and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels respectively; standard errors are in parentheses; random parameter estimation is based on 200 Halton draws.

Figure 4: Cumulative distributions of cost sensitivity across individuals by relative cost position
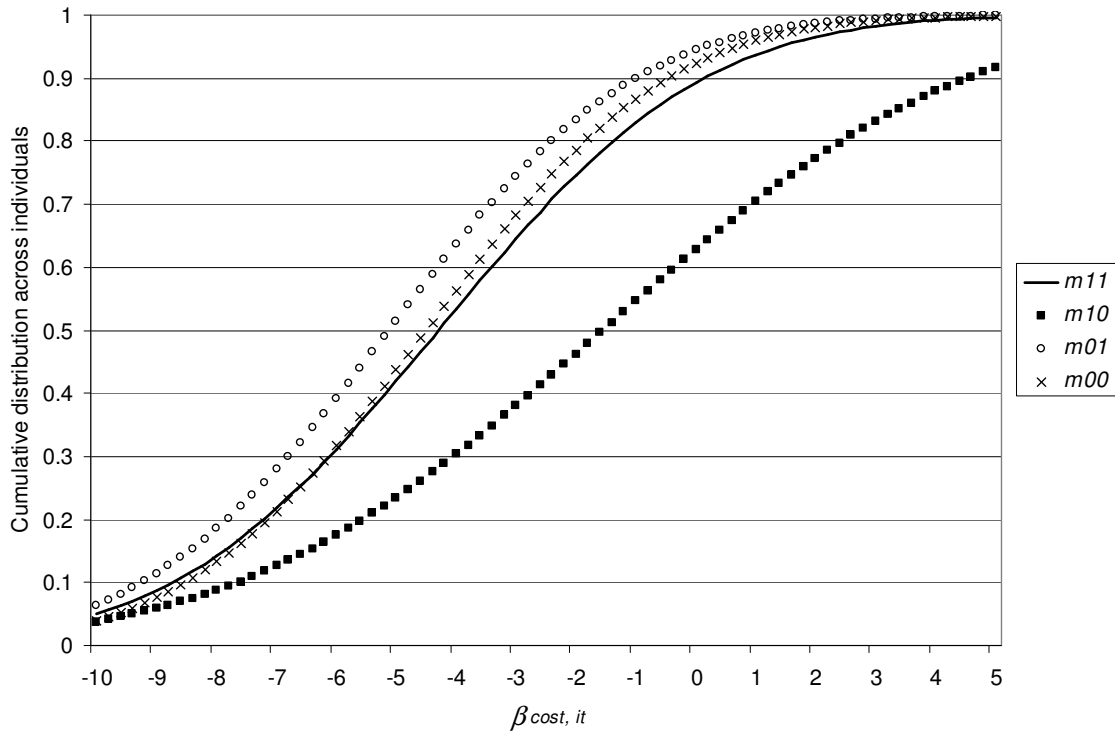
5. Conclusions

The evidence presented in this paper supports the finding of Racevskis and Lupi (2008) that estimates of mean WTP from an elicitation format presenting multiple choice tasks to each respondent are lower than those from an incentive compatible single choice task format. We find no evidence to support the hypothesis that stated preferences in the first choice task presented are affected by advance knowledge that four as opposed to one choice tasks will be presented. In fact, there is equivalence in the evidence on WTP estimates from the first choice task in a sequence and a *SB* choice task. This goes some way to justifying the use of the first choice task in a sequence as an unbiased comparator in studies such as Bateman et al. (2008b) and Day and Pinto (2010). We expect that this

equivalence would be at least as strong where respondents are not informed about how many choice tasks will be presented. This could be investigated by further research.

While there are several possible interpretations for the decrease in WTP after the first question in the *RB* format, our models show that the cost levels presented in previous choice tasks play a role. Three of the behavioural theories in the literature are supported by the evidence. First, the *weak strategic misrepresentation* theory in which respondents might reject an alternative that is preferred to the status quo if a similar good was offered at a lower cost in a previous choice task. Second, the *reference point revision* theory in which value functions shift when a non-status-quo option is chosen. Third, a *cost-driven value learning* theory in which respondents revise their valuation of the good towards a weighted average of the cost levels presented in the sequence to that point (with a higher weight placed on lower cost levels than higher cost levels).

These behavioural explanations have quite different implications. *Weak strategic misrepresentation* implies divergence between stated and true preferences and would lead us to conclude that the property of incentive compatibility is crucial to accurate value estimation. *Reference point revision*, could be accommodated in a repeated choice format using appropriately specified econometric models (Masieroa and Hensher, 2010). The presence of *value learning* has led other authors to conclude that a format comprising a sequence of choice tasks is preferred because it allows respondents to experience the (typically unfamiliar) good (Bateman et al., 2008a). The models herein suggest that values are influenced by the cost levels presented in choice tasks - cost levels that are typically chosen based on statistical efficiency grounds and the anticipated distribution of values over the population, rather than the production cost of the good. This

interdependence raises questions over the appropriate selection of cost levels as well as elicitation format (since presenting multiple choice tasks causes a reduction in value, on average). Our design does not allow further discrimination between the theories (for example, using the scope test applied by (Bateman et al., 2008b)). Regardless, this study confirms the finding of Day and Pinto (2010) that the order in which choice tasks are presented has a significant effect on WTP estimates. An order with increasing cost levels will underestimate WTP relative to an incentive compatible *SB* choice format, while an order with declining cost levels could overestimate WTP.

In this study, similar goods were offered at quite different cost levels over the course of a sequence of choice tasks, making opportunities for strategic response relatively obvious and, potentially, exacerbating *cost-driven value learning* behaviour. We note that Day and Pinto (2010) and Scheufele and Bennett (2010b), in their concurrent and similar study focussing on the case of a pure public environmental good, arrived at similar findings using surveys in which the cost attribute was less dominant. In all of these studies, the cognitive burden of choice tasks was low. Day and Pinto (2010) and Scheufele and Bennett (2010b) employ binary choices between alternatives described by one cost and either one or two non-cost attributes. A key objective for future research will be to establish whether the response behaviour identified in these studies becomes less prevalent as the cognitive burden of the trade-offs in the choice task (potentially measured by the number of attributes attended to and the number of alternatives per choice task) is increased. The body of evidence accumulated thus far casts serious doubt on the standard assumption that all responses are truthful and preferences are stable over the course of a sequence of discrete choice questions.

References

Alfnes F. and Steine G., 2005. None-of-these bias in hypothetical choice experiments. Discussion paper DP-06/05, Department of Economics and Resources Management, Norwegian University of Life Sciences, Aas.

Arrow K., Solow R., Portney P.R., Leamer E.E., Radner R. and Schuman H., 1993. Report of the NOAA panel on contingent valuation. Federal Register 58, 4601-4614.

Baron D. and Myerson R., 1982. Regulating a monopolist with unknown costs. Econometrica 50, 911-930.

Bateman I.J., Burgess D., Hutchinson W.G. and Matthews D.I., 2008a. Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness. Journal of Environmental Economics and Management 55, 127-141.

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

Beenstock M., Goldin E. and Haitovsky Y., 1998. Response bias in conjoint analysis of power outages. Energy Economics 20, 135-156.

Bennett J. and Blamey R., 2001. The Choice Modeling Approach to Environmental Valuation. Edward Elgar, Cheltenham, UK.

Boyle K., Morrison M. and Taylor L., 2004. Why value estimates generated using choice modelling exceed contingent valuation: further experimental evidence. Paper presented at Australian Agricultural and Resource Economics Society Conference, Melbourne, 11-13 February.

Boyle K.J., Bishop R.C. and Welsh M.P., 1985. Starting Point Bias in Contingent Valuation Bidding Games. Land Economics 61, 188-194.

Bradley M. and Daly A., 1994. Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. Transportation 21, 167-184.

Braga J. and Starmer C., 2005. Preference anomolies, preference elicitation and the discovered preference hypothesis. Environmental and Resource Economics 32, 55-89.

Cameron T.A., Poe G.L., Ethier R.G. and Schulze W.D., 2002. Alternative non-market value-elicitation methods: Are the underlying preferences the same? Journal of Environmental Economics and Management 44, 391-425.

Cameron T.A. and Quiggin J., 1994. Estimation using contingent valuation data from a "dichotomous choice with follow-up" questionnaire. Journal of Environmental Economics and Management 27, 218-234.

Carlsson F. and Martinsson P., 2001. Do hypothetical and actual marginal willingness to pay differ in choice experiments? Journal of Environmental Economics and Management 41, 179-192.

____, 2008. Does it matter when a power outage occurs? A choice experiment study on the WTP to avoid power outages. Energy Economics 30, 1232-1245.

Carson K.S., Chilton S.M. and Hutchinson W.G., 2009. Necessary conditions for demand revelation in double referenda. Journal of Environmental Economics and Management 57, 219-225.

Carson R.T. and Groves T., 2007. Incentive and informational properties of preference questions. Environmental and Resource Economics 37, 181-210.

Carson R.T., Groves T. and List J., 2006. Probabilistic influence and supplemental benefits: a field test of the two key assumptions behind using stated preferences. unpublished manuscript.

Carson R.T., Groves T. and Machina M.J., 1997. Stated preference questions: context and optimal response. Paper presented at National science foundation preference elicitation symposium, University of California, Berkeley.

Caussade S., Ortuzar J.d.D., Rizzi L.I. and Hensher D.A., 2005. Assessing the influence of design dimensions on stated choice experiment estimates. Transportation Research Part B 39, 621-640.

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

DCITA, 1998. Putting cables underground. A report for the Minister for Communications, Information Technology and the Arts, November.

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

Farquharson R., 1969. Theory of voting. Yale University Press, New Haven.

Flachaire E. and Hollard G., 2007. Starting point bias and respondent uncertainty in dichotomous choice contingent valuation surveys. Resource and Energy Economics 29, 183-194.

Fosgerau M. and Nielsen S.F., 2006. Deconvoluting preferences and errors: a semi-nonparametric model for binomial panel data. Paper presented at Econometric Society European Meeting, Vienna, Austria, August 24-28.

Green J. and Laffont J.J., 1979. Incentives in public decision making. North-Holland, Amsterdam.

Green J.R. and Laffont J.J., 1978. A sampling approach to the free rider problem, In: Sandmo A (ed), Essays in public economics. Lexington Books, Lexington, MA.

Hanemann W.M., Loomis J. and Kanninen B., 1991. Statistical efficiency of double bounded dichotomous choice contingent valuation. American Journal of Agricultural Economics 73, 1255-1263.

Hensher D.A., 2009. Hypothetical bias, choice experiments and willingness to pay. In Transportation Research Part B.

Hensher D.A. and Greene W.H., 2009. Valuation of travel time savings in WTP and preference space in the presence of taste and scale heterogeneity. Journal of Transport Economics and Policy (in press).

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

Hensher D.A. and Truong T.P., 1985. Valuation of Travel Time Savings: A Direct Experimental Approach. Journal of Transport Economics and Policy 19, 237-261.

Herriges J.A. and Shogren J.F., 1996. Starting point bias in dichotomous choice valuation with follow-up questioning. Journal of Environmental Economics and Management 30, 112-131.

Holmes T. and Boyle K.J., 2005. Dynamic learning and context-dependence in sequential, attribute-based stated-preference valuation questions. Land Economics 81, 114-126.

Hurwicz L., 1972. On informationally decentralized systems, In: McGuire C B and Radner R (ed), Decision and organisation. North-Holland, Amsterdam.

InfraSource Technology, 2007. Undergrounding assessment phase 1 final report. A report for Florida Electric Utilities, February.

IPART, 2002. Electricity Undergrounding in New South Wales. A report for the Minister for Energy, May.

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

Ladenburg J. and Olsen S.B., 2008. Gender-specific starting point bias in choice experiments: Evidence
        from an empirical study. Journal of Environmental Economics and Management 56, 275-285.

Loeb M. and Magat W., 1979. A decentralized method for utility regulation. Journal of Law and Economics
        22, 399-404.

Louviere J.J. and Woodworth G., 1983. Design and analysis of simulated consumer choice or allocation
        experiments: an approach based on aggregate data. Journal of Marketing Research 20, 350-367.

Lusk J. and Schroeder T., 2004. Are choice experiments incentive compatible? A test with quality
        differentiated beef steaks. American Journal of Agricultural Economics 86, 467-482.

Masieroa L. and Hensher D.A., 2010. Shift of reference point and implications on behavioral reaction to
        gains and losses. unpublished manuscript.

McFadden D., 1980. Econometric models for probabilistic choice among products. Journal of Business 53,
        13-29.

McFadden D. and Leonard G., 1995. Issues in the contingent valuation of environmental goods:
        Methodologies for data collection and analysis, In: Hausman J A (ed), Contingent Valuation: A
        Critical Assessment. North-Holland, Amsterdam.

Mirrlees J., 1971. An exploration in the theory of optimal income taxation. Review of Economic Studies
        38, 175-208.

NERA Economic Consulting and ACNielsen, 2003. Willingness to pay research study. A report for
        ACTEW Corporation and ActewAGL, September.

Plott C.R., 1996. Rational individual behavior in markets and social choice processes: the discovered
        preference hypothesis, In: Arrow K, Colombatto E, Perleman M and Schmidt C (ed), Rational
        foundations of economic behavior. Macmillan, London.

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

Rose J.M., Hess S., Bliemer M.C.J. and Daly A., 2009. The impact of varying the number of repeated

    choice observations on the mixed multinomial logit model. Paper presented at European Transport

    Conference, Leeuwenhorst, The Netherlands, 5-7 October.

Samuelson P.A., 1954. The pure theory of public expenditure. Review of Economics and Statistics 36, 387-

    389.

Scarpa R. and Rose J., 2008. Design Efficiency for Non-Market Valuation with Choice Modelling: How to

    Measure it, What to Report and Why. Australian Journal of Agricultural and Resource Economics

    52, 253-282.

Scheufele G. and Bennett J., 2010a. Effects of alternative elicitation formats in discrete choice experiments.

    Paper presented at 54th annual Australian Agricultural and Resource Economics Society

    conference, Adelaide, South Australia, 9-12 February.

____, 2010b. Ordering effects and response strategies in discrete choice experiments. Paper presented at

    World Congress of Environmental and Resource Economists, Montreal, Canada, 28 June - 2 July.

Schleifer A., 1985. A theory of yardstick competition. Rand Journal of Economics 16, 319-327.

Small K.A. and Rosen H.S., 1981. Applied Welfare Economics with Discrete Choice Models.

    Econometrica 49, 105-130.

Spence M.A., 1975. Monopoly, quality, and regulation. The Bell Journal of Economics 6, 417-429.

List of previous EMD Occasional Papers

Occasional Paper No. 1
A Comparison of Contingent Valuation and Choice Modelling: estimating the environmental values of Catalonian Forests,
Joan Mogas, Pere Riera and Jeff Bennett
December 2002

Occasional Paper No. 2
Chinese Agricultural Water Resource Utilisation in the 21st Century,
Han Hongyun and Jeff Bennett
January 2003

Occasional Paper No. 3
Estimating the Costs of Atmospheric Carbon Reductions in Mexico,
Paola Del Rio and Jeff Bennett
February 2003

Occasional Paper No. 4
Transfer of choice model benefits: a case study of stream mitigation,
Geoffrey N. Kerr and Basil M.H. Sharp
December 2003

Occasional Paper No. 5
Financial viability of forest certification in industrial plantations: a case study from the Solomon Islands ,
Fabio Pesce and Padma Lal
March 2004

Occasional Paper No. 6
Analysis of Learning Cycles in Participatory Environment and Development Projects: Lessons from Nepal,
Chiharu Hiyama and Meg Keen
September 2004

Occasional Paper No. 7
Transaction costs and the assessment of greenhouse policies in the transport energy sector,
Albert Ofei-Mensah and Jeff Bennett
November 2004

Occasional Paper No. 8
An Economic Valuation of Wetlands in Vietnam's Mekong Delta: a case study of direct use values in Camau Province,
Thang Nam Do and Jeff Bennett
June 2005

[**Occasional Paper No. 9**](#)
On the Theory of Decentralization, Forests and Livelihoods,
Luca Tacconi, Yulia Siagian, Ronny Syam
April 2006

[**Occasional Paper No. 10**](#)
Valuing Australian Cultural Institutions: Developing a Cultural Worldview Scale
Andy Sungnok Choi, Franco Papndrea, Jeff Benntt
October 2006

[**Occasional Paper No. 11**](#)
Forests, Agriculture, Poverty and Land Reform: The Case of the Indonesian Outer Islands
Luca Tacconi and Iwan Kurniawan
November 2006

[**Occasional Paper No. 12**](#)
Consumer Demand for Sustainable Wild-caught and Cultured Live Reef Food Fish in Hong Kong
Noel Wai Wah Chan, Jeff Bennett, Brian Johnston
December 2006

[**Occasional Paper No. 13**](#)
House prices and underground electricity distribution lines: the case of three selected suburbs in Canberra
Ben McNair
February 2009