Journal of Choice Modelling, 5(1), 2012, pp 77-97
www.jocm.org.uk

Are there specific design elements of choice experiments and types of people that influence choice response certainty?

David A. Hensher[1, †], John M. Rose[1, †] and Matthew J. Beck[1,Ŧ]

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

Received 16 May 2011, revised version received 13 September 2011, accepted 2 April 2012

Abstract

The development of more realistic choice experiments has taken on board a number of suggestions in the broader hypothetical bias literature. One issue, in particular, is the increasing interest in finding ways to bridge the gap between the stated choice response and real choosing, as a way of increasing the confidence with which an individual would hypothetically purchase or use an alternative that is actually chosen in the choice experiment. In this paper we investigate the relationship between the respondent's response to a certainty question, defined on a 1-10 scale of surety, and features of the choice experiment that may have a bearing on the degree of confidence that can be placed on the stated choice, controlling for exogenous effects such as socioeconomic characteristics and attitudes to vehicle emissions. The focus on response certainty in this paper is as an external validity test. We find, using a generalised ordered logit model, compelling evidence that the number of acceptable alternatives and hence associated levels of attributes, together with the contrast of attribute levels of each designed alternative relative to an experienced status quo (or reference) alternative, play an important role in establishing certainty of response in a real market. The evidence should be taken on board in the future design of more realistic choice experiments.

Keywords: Response certainty, ordered choice, random thresholds, automobile choice, stated choice experiment, internet survey, context influences, reference alternative

* Corresponding author, T: +61-2-93510071, F: + 61-293510088, david.hensher@sydney.edu.au
† T: +61-2-93510168, F: + 61-293510088, john.rose@sydney.edu.au
Ŧ T: 61-2-93510077, F: + 61-293510088, matthe.beck@sydney.edu.au

ISSN 1755-5345

1 Introduction

Choice studies can be characterised by three key elements – attributes, alternatives, and choice responses. In recent years, an increasing number of analysts have highlighted a concern with the assumption, in the majority of choice studies, that all attributes are traded in a fully compensatory manner and are by implication all relevant, and that each attribute and its trade is treated by the individual decision maker as totally certain (see e.g., Swait 2001, Cantillo et al. 2006, Hensher and Collins 2011). There is now a burgeoning literature on attribute processing (see Leung and Hensher 2011 for an overview). What has been given much less attention is the extent to which a respondent is certain of actually choosing the alternative that they indicated was their preferred alternative in a designed choice set if it were offered in a real market. Certainty response is one possible way of accounting for the risk that one might attach to the choice of an alternative in a choice experiment. Certainty discussed below in the context of consistency between hypothetical and real choices, is by no means the only way that certainty could be interpreted, and depends to a large extent on how the certainty question is phrased. We focus on one interpretation, namely the role of response certainty as a guide to the external validity of the stated choice response. Efforts to understand what are some key drivers underlying the design of the choice experiment that influence respondent choice response and hence certainty of such a choice being made if it were offered in a real market is growing with a small literature, predominantly focussed on public goods. This paper is a contribution to this literature, exploring candidate influences on choice certainty in a private good context using advanced developments in ordered choice modelling.

The hypothetical bias literature in particular has focussed on the certainty of response associated with a choice experiment if the alternative were offered in a real market. Johannesson et al. (1999), Fuji and Garling (2003) and Lundhede et al. (2009) offer some ideas on the certainty scale. Supplementary questions are increasingly being included to establish '*the confidence with which an individual would hypothetically purchase or use the good (or alternative) that is actually chosen in the choice experiment'*; the latter being added into the choice experiment after each choice scenario. Johannesson et al. (1999) proposed a supplementary *certainty scale* question after each choice scenario, on a scale 0 (very unsure) to 10 (very sure), to indicate how sure or certain the respondent is that they would actually choose a particular alternative (or not at all) at the indicated attribute levels. This response metric can be used to exogenously weight the sample to represent a way of placing a higher weight on those choices that one has more confidence in actually being made[1].

Li and Mattson (1995) and Loomis and Ekstrand (1998) have suggested ways to incorporate uncertainty for all responses directly in the likelihood function, through the uncertainty level stated on a scale post decision, along similar lines to Johannesson et al. (1999). Loomis and Ekstrand (1998) found a statistically significant link between familiarity with a good (the spotted owl),

[1] An interesting way of including response certainty into a model is to create a relative measure around a reference alternative, where the latter has been chosen in a real market and hence its certainty value is 10 on the 1-10 scale. Deviations from 10 may be more informative than the actual certainty scale value.

measured as low (1), sample average (1.9) and high (3), experimental bid level (range of US$1 to US$350), and certainty in the rating of the respondent's certainty in a WTP answer in a contingent valuation experiment. Brouwer et al. (2010) added an additional question after each choice task to identify how certain respondents felt about their choices. The responses were identified on a semi-itemised 0–10 rating scale, where 0 means completely uncertain and 10 completely certain. They wanted to see whether respondents felt they became more confident and hence certain as a result of experience and learning as they went through the choice sequence. Self-reported choice certainty from the survey was regressed on a number of possible explanatory factors in an ordered probit model.

Like Lundhede et al. (2008), they were also interested to provide further empirical evidence of the hypothesis that choice uncertainty increases as alternatives become less distinguishable from each other in terms of the utility they generate. Regressing on declared certainty, they find evidence of the role of respondent socioeconomic characteristics; for example, women as well as older individuals are less certain, and higher income individuals are more certain; as well the choice features have an influence, especially a higher difference in utility between alternatives which results in more certainty.

Olsen et al. (2011) also investigated the role of a number of socioeconomic characteristics, but the most informative finding supports the evidence in Brouwer et al. (2010) that the utility difference is associated with more certainty. Bech et al. (2011), using a sample of 1053 respondents exposed to 5, 9 or 17 choice sets in a discrete choice experiment designed to elicit preferences for dental services in Denmark, found no differences in response rates and no systematic differences in the respondents' self reported perception of the uncertainty of their discrete choice experiment answers. Champ and Bishop (2001) developed a contingent valuation experiment to estimating actual willingness to donate based on contingent donations, with a follow-up question in which respondents rate on a ten point scale (with endpoints labelled 1 = very uncertain and 10 = very certain), how certain they were that they would purchase (or not purchase) wind generated electricity that was offered to them. An ordered probit model with the level of certainty as the dependent variable provided insights into possible sources of uncertainty. They found that the level of certainty is not related in a statistically significant manner to the offer amount, in contrast to Samnaliev et al. (2006) who found the opposite in a study of attitudes towards user fees to access public lands in the context of the current US Fee Demonstration Program (FDP). A higher certainty level was expressed by respondents who agreed or strongly agreed that the program is worth the extra cost, liked the idea of wind-generated electricity, and frequently donated money to environmental causes. Respondents who agreed or strongly agreed with the statements "I can't afford to pay the extra cost of wind generated electricity" and "Electricity costs are too high" were less certain about their positive response to the contingent donation question. These results seem consistent with the argument that some respondents who say *yes* to the contingent donation question are expressing a positive sentiment toward the program but not specifically agreeing to pay the offer amount.

Lundhede et al. (2009) focus on the evaluation of different approaches to using respondents' stated certainty in choice to improve model estimation, noting that "how the researcher handles respondents' stated certainty will depend on what is assumed to be the reasons for the stated certainty" (page 120). They cite Samnaliev et al. (2006) who summarise four assumptions or hypotheses (in the context of contingent valuation studies): (i) certainty levels indicated by respondents will reflect their attempt to appear consistent in answers, (ii) certainty levels may be susceptible to protesting and strategic behaviour such as respondents exaggerating certainty, (ii) in the context of preference uncertainty, respondents use stated uncertainty to scale down their stated willingness to pay (WTP), and (iv) respondents are rational, truth-telling and non-strategic, but may assess the value of a change with some degree of uncertainty and, therefore, the response itself may be subject to error which translates into a probability that the respondent does not choose the utility maximising alternative.

Our interest is in the fourth interpretation, where we suggest that the choice of interest, a private good (i.e., automobile fuel type choice), is most likely to be non-strategic, or at least far less strategic than environmental (public good) and political applications, especially when the latter studies predominantly use contingent valuation instead of a choice experiment. It is well known that CV studies run the risk of strategic response, in contrast to choice experiments. We acknowledge that the environmental impact of different auto vehicles fuel types may cause some of the same issues to occur as in public goods evaluation (strategic choice, uncertainty over actual impacts, etc.), even if to a far lesser extent; however we suggest that the choice of a vehicle type is very likely to be consistent with utility maximisation even when 'green choices' matter to specific individuals.

This paper investigates the influences on choice response certainty in a choice experiment with multiple attributes in the context of automobile purchase preferences (a private good) in Sydney, involving respondents choosing amongst petrol, diesel and hybrid fuelled vehicles (associated with specific levels of fuel efficiency and engine capacity) when faced with a mix of vehicle prices, fuel prices, fixed annual registration fees, annual emission surcharges, and vehicle kilometre emission surcharges. The focus in the current paper is not on estimating a vehicle fuel type choice model (see Hensher et al. 2011 and Beck et al. 2011 for such models), but on studying the relationship between the respondent's response to a certainty question, defined on a 1-10 scale of surety, and features of the choice experiment that may have a bearing on the degree of confidence we can place on the stated choice, controlling for exogenous effects such as socioeconomic characteristics and attitudes to vehicle emissions. We use a generalised ordered logit model to obtain evidence on systematic sources of variation in choice certainty.

The paper is organised as follows. We first set out the ordered choice model specification that defines the certainty scale as the dependent variable. The data is then presented, followed by model estimation results and interpretation.

2 Response Certainty as an Ordered Choice

The response certainty scale used herein is a 10 point scale with a natural ordering from 1 (very unsure) to 10 (very sure). Our interest is in identifying systematic influences on variations in a sample's response along this scale, recognising that we have selected 10 points on an underlying continuous distribution. The cut-off levels on the scale are likely to be perceived differently by each sampled respondent (and even possibly between the choice sets that each person assesses), suggesting that fixed cut-offs or thresholds that fail to account for preference heterogeneity, both random and systematic, may be inappropriate.

Furthermore the real possibility exists for different variances in unobserved effects (or heteroscedasticity) defined through the random error component of the utility expression, including the possibility of systematic sources of influence. This set of candidate sources of explanation of differences in subjective response choice certainty can be tested for in a generalised ordered choice model, of the logit form, that has been developed by Greene and Hensher (2010) with the key elements summarised below. The approach set out below is a behaviourally richer representation of ordered choice, extending beyond simple ordered logit and probit (with fixed thresholds and preference homogeneity in attribute parameter estimates), the method used in existing choice certainty studies where an ordered choice model is selected[2].

The model foundation is an underlying random utility or latent regression model, of the form in equation (1) in which the continuous latent utility, y_i^* is observed in discrete form through a censoring mechanism (equation 2).

$$y_i^* = \boldsymbol{\beta}'\mathbf{x}_i + \varepsilon_i, \tag{1}$$

where

$$\begin{aligned} y_i &= 0 \text{ if } \mu_{-1} < y_i^* \leq \mu_0, \\ &= 1 \text{ if } \mu_0 < y_i^* \leq \mu_1, \\ &= 2 \text{ if } \mu_1 < y_i^* \leq \mu_2 \\ &= \ldots \\ &= J \text{ if } \mu_{J-1} < y_i^* \leq \mu_J. \end{aligned} \tag{2}$$

The model contains the unknown marginal utilities, $\boldsymbol{\beta}$, as well as J+2 unknown threshold parameters, μ_j, all estimated using a sample of n observations, indexed by $i = 1,\ldots,n$. The data consist of the explanatory variables, $\mathbf{x}_i$ and the observed discrete outcome (or certainty scale), $y_i = 0,1,\ldots,J$. The disturbance term, ε_i is continuous with cumulative distribution function, $F(\varepsilon_i|\mathbf{x}_i) = F(\varepsilon_i)$ and with density $f(\varepsilon_i) = F'(\varepsilon_i)$. The assumption of the distribution of ε_i includes independence from $\mathbf{x}_i$. The probabilities associated with the observed outcomes are given as equation (3).

$$\text{Prob}[y_i = j \mid \mathbf{x}_i] = \text{Prob}[\varepsilon_i \leq \mu_j - \boldsymbol{\beta}'\mathbf{x}_i] - \text{Prob}[[\varepsilon_i \leq \mu_{j-1} - \boldsymbol{\beta}'\mathbf{x}_i], j = 0,1,\ldots,J. \tag{3}$$

The identifying restriction σ_ε = a known constant, $\bar{\sigma}$, is imposed and it is assumed that $\text{Var}[\varepsilon_i|\mathbf{x}_i] = \pi^2/3$ in the logit model. The likelihood function for estimation of the model parameters is based on the implied probabilities given in equation (4)[3].

[2] In addition, the traditional ordered choice model fails to take into account the panel nature of data that is common in stated choice experiments.

[3] Several normalisations are needed to identify the model parameters. First, given the continuity assumption, in order to preserve the positive signs of the probabilities, we require $\mu_j > \mu_{j-1}$. Second, if the support is to be the entire real line, then $\mu_{-1} = -\infty$ and $\mu_J = +\infty$. Finally, assuming that $\mathbf{x}_i$ contains a constant term, we will require $\mu_0 = 0$. With a constant term present, if this normalisation is not imposed, then adding any nonzero

$$\text{Prob}[y_i = j \mid \mathbf{x}_i] = F(\mu_j - \boldsymbol{\beta}'\mathbf{x}_i) - F(\mu_{j-1} - \boldsymbol{\beta}'\mathbf{x}_i) > 0, j = 0,1,...,J. \quad (4)$$

The basic model (4) can be enhanced to allow for a number of ways in which individual preference heterogeneity can be accounted for in the marginal utilities, in the threshold parameters, and in the scaling (variance) of the **random** components. The intrinsic heterogeneity in utility functions across individuals is captured by writing:

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\Delta}\mathbf{z}_i + \boldsymbol{\Gamma}\mathbf{v}_i \quad (5)$$

where $\boldsymbol{\Gamma}$ is a lower triangular matrix and $\mathbf{v}_i \sim \text{N}[\mathbf{0},\mathbf{I}]$, and $\mathbf{z}_i$ is a set of observed individual-specific influences on marginal utility. $\boldsymbol{\beta}_i$ is normally distributed across individuals with conditional mean given in equation (6).

$$E[\boldsymbol{\beta}_i|\mathbf{x}_i,\mathbf{z}_i] = \boldsymbol{\beta} + \boldsymbol{\Delta}\mathbf{z}_i \quad (6)$$

and conditional variance in equation (7).

$$\text{Var}[\boldsymbol{\beta}_i|\mathbf{x}_i,\mathbf{z}_i] = \boldsymbol{\Gamma}\mathbf{I}\boldsymbol{\Gamma}' = \boldsymbol{\Omega}. \quad (7)$$

This is a random parameters formulation. The thresholds are also modelled randomly and nonlinearly as:

$$\mu_{ij} = \mu_{i,j-1} + \exp(\alpha_j + \boldsymbol{\delta}'\mathbf{r}_i + \sigma_j w_{ij}),\ w_{ij} \sim \text{N}[0,1] \quad (8)$$

with normalisations and restrictions $\mu_{-1} = -\infty$, $\mu_0 = 0$, $\mu_J = +\infty$. For the remaining thresholds, we have equation system (9).

$$\begin{aligned}
\mu_1 &= \exp(\alpha_1 + \boldsymbol{\delta}'\mathbf{r}_i + \sigma_1 w_{j1}) \\
&= \exp(\boldsymbol{\delta}'\mathbf{r}_i)\exp(\alpha_1 + \sigma_1 w_{j1}) \\
\mu_2 &= \exp(\boldsymbol{\delta}'\mathbf{r}_i)\,[\exp(\alpha_1 + \sigma_1 w_{j1}) + \exp(\alpha_2 + \sigma_2 w_{j2})], \\
\mu_j &= \exp(\boldsymbol{\delta}'\mathbf{r}_i)\left(\Sigma_{m=1}^{j}\exp(\alpha_m + \sigma_m w_{im})\right), j = 1,...,J\text{-}1 \\
\mu_J &= +\infty.
\end{aligned} \quad (9)$$

This formulation preserves the ordering of the thresholds and incorporates the necessary normalisations. It also allows observed variables and unobserved heterogeneity to play a role both in the utility function and in the thresholds. The thresholds, like the regression itself, are shifted by both observable ($\mathbf{r}_i$) (which could contain the same covariates as $\mathbf{z}_i$) and unobserved (w_{ij}) heterogeneity. The model is fully consistent, in that the probabilities are all positive and sum to one

constant to μ_0 and the same constant to the intercept term in $\boldsymbol{\beta}$ will leave the probability unchanged. Given the assumption of an overall constant, only J-1 threshold parameters are needed to partition the real line into the J+1 distinct intervals. The identification issues associated with unordered choice models in selecting the relevant base alternative-specific constant (as so eloquently shown in Joan Walker's research – see Chiou and Walker 2007), is not an issue in ordered choice models.

by construction. The disturbance variance is allowed to be heteroscedastic, specified randomly as well as deterministically. Thus,

$$\mathrm{Var}[\varepsilon_i|\mathbf{h}_i,e_i] = \sigma_i^2 = \exp(\boldsymbol{\gamma}'\mathbf{h}_i + \tau e_i)^2 \tag{10}$$

where $e_i \sim N[0,1]$, $\mathbf{h}_i$ are observed covariates and $\boldsymbol{\gamma}'$ are estimates parameters, Define $\mathbf{v}_i = (v_{i1},...,v_{iK})'$ for K attributes and $\mathbf{w}_i = (w_{i1},...,w_{i,J-1})'$. Combining terms, the conditional probability of outcome is given in equation (11).

$$\mathrm{Prob}[y_i = j \mid \mathbf{x}_i,\mathbf{z}_i,\mathbf{h}_i,\mathbf{r}_i,\mathbf{v}_i,\mathbf{w}_i,e_i] = F\left[\frac{\mu_{ij} - \boldsymbol{\beta}_i'\mathbf{x}_i}{\exp(\boldsymbol{\gamma}'\mathbf{h}_i + \tau e_i)}\right] - F\left[\frac{\mu_{i,j-1} - \boldsymbol{\beta}_i'\mathbf{x}_i}{\exp(\boldsymbol{\gamma}'\mathbf{h}_i + \tau e_i)}\right] \tag{11}$$

The model is estimated by maximum simulated likelihood (see details in Greene and Hensher 2010).

3 The Choice Experiment and Survey Process

We draw on a choice experiment that was designed for a study whose main objective was to identify possible ways to reduce emissions from automobile ownership and use. Each choice scenario was accompanied by a **s**upplementary question on the certainty that the respondent would actually make that choice (Figure 1).

The labelled choice experiment was defined on three fuel type alternatives - petrol, diesel and hybrid. Within each fuel class, each alternative was further defined by a vehicle class: small, luxury small, medium, luxury medium, large and luxury large, to ensure that the experiment would have adequate attribute variance as well as meaningful attribute levels over the alternatives, particularly with respect to price, whilst still having a manageable number of alternatives for the design. Nine attributes were included in the choice experiment, refined via review of the available literature on vehicle purchasing, as well as through a pilot survey (Beck et al. 2009) and preliminary analysis of secondary data sets. The attributes and their levels are summarised in Table 1.

Tables A1 and A2 in the appendix show the levels chosen for the annual and variable surcharges. Both of the surcharges are determined by the type of fuel a vehicle use**s** and the fuel efficiency of that vehicle. For a given vehicle, if it is fuelled by petrol, owners would pay a higher surcharge than if it was fuelled by diesel, which is in turn more expensive than if it was a hybrid. Once the car has been specified in terms of fuel type and efficiency, there are five levels of surcharge that could be applied.

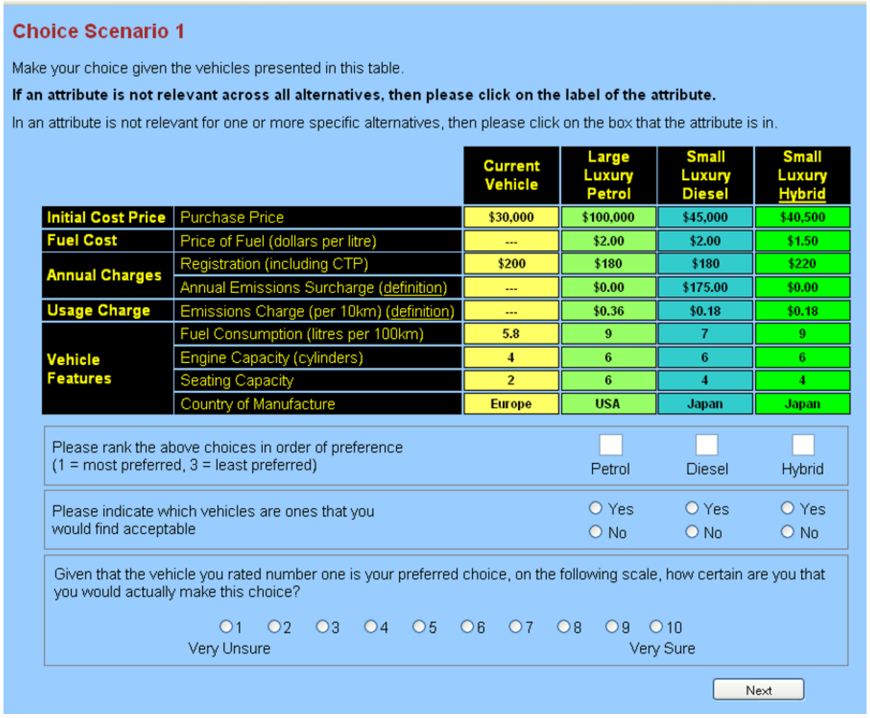

Figure 1 Illustrative Stated Choice Screen

Table 1 Attribute Levels for Choice Experiment

	Levels	1	2	3	4	5
Purchase Price	*Small*	$15,000	$18,750	$22,500	$26,250	$30,000
	Small Luxury	$30,000	$33,750	$37,500	$41,250	$45,000
	Medium	$30,000	$35,000	$40,000	$45,000	$50,000
	Medium Luxury	$70,000	$77,500	$85,000	$92,500	$100,000
	Large	$40,000	$47,500	$55,000	$62,500	$70,000
	Large Luxury	$90,000	$100,000	$110,000	$120,000	$130,000
Fuel Price ($/litre)	*Pivot off daily price*	-25%	-10%	0%	10%	25%
Registration (incl. CTP)	*Pivot off actual purchase*	-25%	-10%	0%	10%	25%
Annual Emissions Charge	*Pivot off fuel efficiency*	Random allocation of one of five levels (see Appendix Table A1)				
Variable Emissions Charge	*Pivot off fuel efficiency*	Random allocation of one of five levels (see Appendix Table A2)				
Fuel Efficiency (ltr/100km)	*Small*	6	7	8	9	10
	Medium	7	9	11	13	15
	Large	7	9	11	13	15
Engine Cylinders	*Small*	4	6			
	Medium	4	6			
	Large	6	8			
Seating Capacity	*Small*	2	4			
	Medium	4	5			
	Large	5	6			

The choice experiment is a D-efficient design where the focus is on the asymptotic properties of the standard errors of estimates, given the priors of attribute parameters. Prior parameter estimates obtained from substantive pilot surveys are used to minimise the asymptotic variance-covariance matrix which leads to lower standard errors and more reliable parameter estimates, for a given sample size (see Rose and Bliemer 2008 for details). The methodology focuses not only on the design attributes which are expanded out through treatment repetition, i.e., multiple choice sets, but also on the non-expanded socio-demographics and other contextual variables that are replicated as constants within each observation, and whose inclusion should have the greater influence on the efficient sample size.

A reference alternative is identified prior to the choice scenarios and it describes a recent purchase of a car (in the period 2007 to 2009). The reference alternative acts as a pivot for the experimental design for the known attributes of the alternative (see Rose et al. 2008). For the petrol, diesel and hybrid alternatives, all attributes vary, and the combinations of levels are optimised via the design process. The size of each vehicle for each fuel type alternative varies randomly and is endogenous to the design. The level of the annual and variable surcharge that appears in each alternative is conditional on the fuel type and efficiency of the vehicle. The values of fuel price and registration (including compulsory third party (CTP) insurance) pivot off an actual reference alternative as follows:

- Fuel price pivots around the daily fuel price as entered by the interviewer. There are five levels of fuel price (-25%, -10%, no change, +10%, +25%).
- Registration (including CTP) pivots around the actual cost provided by the respondent. There are five levels of registration (-25%, -10%, no change, +10%, +25%).
- The annual emissions surcharge is determined by the type of fuel used by the alternative and the fuel efficiency of that vehicle. For each fuel type and fuel efficiency combination, there are five levels of surcharge that apply (Table A1).
- The variable emissions surcharge is determined by the type of fuel used by the alternative and the fuel efficiency of that vehicle. For each fuel type and fuel efficiency combination, there are five levels of surcharge that apply (Table A2).

An internet based survey with face to face assistance of an interviewer was programmed. An eligible respondent had to have purchased a new vehicle in 2007, 2008 or 2009. Details of response rates and reasons for non-eligibility are summarised in Beck et al. (2011). The survey was completed online at a central location (varied throughout the Sydney metropolitan area to minimise travel distance for respondents). Respondents provided details of the vehicles within the household, and details of the most recent (or a potential) purchase. A number of choice sets are provided (with an example shown in Figure 1), with all participants asked to review the alternatives, and then indicate their preferred outcome, as well as an indication of which alternatives are acceptable, and what is the certainty of actually making the choice if it were available now in a real market. The response certainty scale used herein is a 10 point scale with a natural ordering from 1 (very unsure) to 10 (very sure). It is common in modelling to recode between ranking and choice, which is required in order to indicate the degree of choice certainty; however, the idea that a respondent can easily do this should not be assumed. Although there may be a critical mental shift required from the respondent to conciliate the ranking (equal to one) with a certainty of choice indication, we have assumed that respondents are equally capable of doing this translation compared to a situation where only the chosen is requested.

A series of attitudinal question were also asked, and shown in Figure 2. Further details of the overall study are given in Beck et al. (2011).

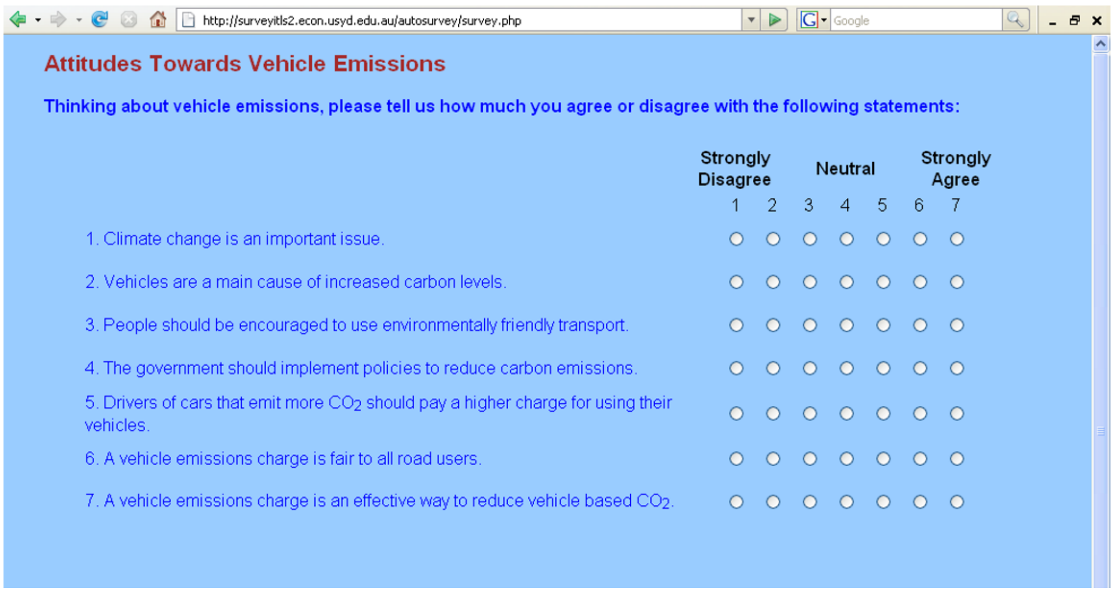

Figure 2 Attitudinal Questions

4. Empirical Results

The data was collected over a four month period in 2009. The final sample used in model estimation comprises 5,700 choice sets, a subset of the full data set. The data contained respondents who completed eight choice sets. Given the focus in this paper is on the role of choice response certainty, we refer readers to Hensher et al. (2011) and Beck et al. (2011) for details of the fuller data set, confining the presentation to the data elements relevant to the modelling undertaken below. Table 2 summarises the ordered choice models, including the mean of each data items used in the estimation of the models.

The choice certainty dependent variable was transformed into seven levels, with levels 1-4 combined given that there were so few responses in individual levels in this range. The overall mean across all eight choice scenarios and respondents is 7.289 with a standard deviation of 2.075. Although 7.289 suggests a relatively high degree of surety, there are about 28 percent of choice set situations where the degree of certainty is below 5. For the combined sample, the frequency of each level of response certainty is 9.81 percent for levels 1-4 (unsure), 6.91 percent for level 5, 11.33 percent for level 6, 18.53 percent for level 7, 23.96 percent for level 8, 16.23 percent for level 9 and 13.23 percent for levels 10 (very sure). We also report the certainty responses for each of the eight choice sets in a sequence. There is very little variation across choice sets; the mean and standard deviation (in brackets) for choice sequences 1 to 8 are respectively 7.35 (2.01), 7.26 (2.06), 7.19 (2.13), 7.33 (2.04), 7.38 (2.20), 7.37 (2.11), 7.27 (2.22), and 7.33 (2.17). Thus, at aggregate sample level, we would conclude that there is no evidence of choice sequence bias in choice certainty response.

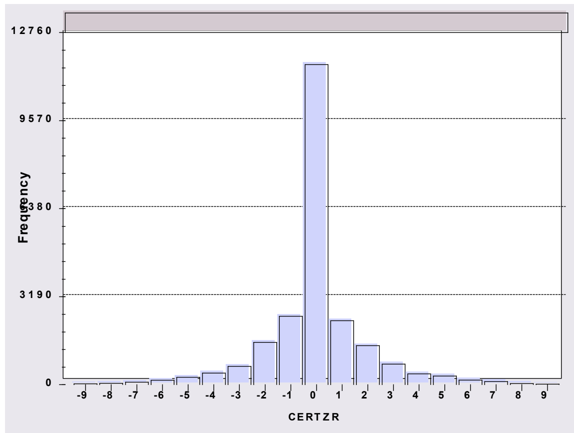

Figure 3 Distribution Range of Response Certainty across Eight Choice Sequences at the Respondent Level

We also derived the difference in response certainty at a respondent level across the eight choice sets in sequence and then averaged the evidence at the individual respondent level. A plot of the evidence is given in Figure 3. We find very small differences, typically plus or minus 2 ratings on the scale across the sample, with 51 percent being identical, 20 percent within a single level difference, and 13 percent within a 2 level difference, a total of 85 percent. What this suggests is that if we recognise the possibility of bands of possible similarity as surety, then we can conclude that respondents appear to exhibit very similar response certainty behaviour across all 8 choice sets in the sequence. This could either mean that they simply indicated the same response without thinking (although we doubt this given the evidence in the model), or that choices on offer were such that they were accepting of a reasonably wide range of offerings as a set from which they would actually choose any of them if available in a real market.

The final generalised ordered choice model is estimated using 500 Halton draws, and accounts for the panel nature of the data (i.e., eight observations per individual)[4]. A simple ordered logit model is also estimated and summarised in Table 2 as a basis of comparison; however this model form is not able to account for the panel nature of the data, and nor can it allow for random thresholds and random heteroscedasticity, although it can allow for systematic decomposition of the error variance.

The overall goodness of fit of the ordered logit model is very poor (i.e., -10315.75, a ρ^2 of 0.036) compared to the generalised ordered logit (GoL) model (i.e., -7930.23 and ρ^2 of 0.231). It is clear that the focus must be on the GoL model, which we now consider in some detail. The generalised ordered logit model contains three points at which changes in the observed variables can induce changes in the probabilities of the outcomes: in the thresholds, μ_{ij}, in the marginal utilities, $\boldsymbol{\beta}_i$, and in the utility function, $\mathbf{x}_i$.

[4]We ran GOL models with 100, 250, 400, 500 and 1000 draws and found that the parameter estimates stabilised at around 500 draws.

Table 2 Summary of Model Results

Note: GOL took over 4 hours to estimate

Explanatory variables	**Mean**	**Ordered Logit**		**Generalised Ordered Logit**	
		Par.	(*t*-ratio)	Par.	(*t*-ratio)
Constant	-	1.3917	(16.90)	1.4829	(6.68)
Gender (male=1)	0.487	0.1077	(5.09)	0.0197	(0.30)
Number of acceptable alternatives per choice set	2.59	-0.0112	(-1.03)	0.3128	(7.13)
Number of persons involved in recent purchase decision	2.01	-0.1222	(-7.09)	0.0119	(0.25)
Australian manufactured vehicle (1,0)	0.367	0.0587	(3.01)	0.3305	(4.37)
Personal income ($000s)	74.26	-0.0007	(-2.16)	0.0009	(0.73)
Number years held driver's licence	26.40	-0.0046	(-5.54)	0.0079	(2.27)
Full time employed (1,0)	0.586	0.1255	(4.27)	0.2074	(1.93)
Part time employed (1,0)	0.192	-0.0468	(-1.58)	-0.1441	(-1.64)
Climate change important issue	5.460	-0.0372	(-4.60)	0.3831	(11.60)
People should be encouraged to use environmentally friend transport	5.706	0.0260	(2.82)	-0.0745	(-2.67)
Govt. should implement carbon reduction policies	5.503	0.0158	(1.95)	0.0779	(2.95)
Drivers of high CO_2 cars should pay more	4.542	-0.0142	(-1.90)	0.1228	(4.55)
Vehicle emissions charge is fair to all road users	4.065	-0.0343	(-4.62)	-0.0748	(-3.19)
A vehicle emissions charge is effective way to reduce vehicle based CO2	3.964	0.0432	(6.17)	-0.1982	(-7.86)
Household income ($000s)	115.83	-0.0006	(-3.67)	-0.0137	(-16.90)
Number of children in household	0.79	-0.0239	(-3.05)	-0.3350	(-11.70)
Number of adults in household	2.14	-0.0228	(-1.22)	0.1068	(1.55)
Number of household members with a drivers licence	2.10	0.0469	(2.51)	0.2284	(3.23)
Fuel cost (SC level < status quo level) (negative) Hybrid	-0.1948	0.0978	(2.05)	0.0416	(0.55)
Regn cost (SC level > status quo level) (positive) Petrol	115.11	-0.0002	(-3.62)	-0.0020	(-15.8)
Regn cost (SC level < status quo level) (negative) Petrol	-105.08	0.00005	(0.66)	-0.0027	(-15.7)
Regn cost (SC level < status quo level) (negative) Hybrid	-104.74	0.0003	(3.38)	0.0026	(16.10)
Fuel efficiency (SC level > status quo level) (positive) Diesel	2.31	-0.0006	(-0.14)	0.0170	(2.12)
Fuel efficiency (SC level < status quo level) (negative) Diesel	-2.38	-0.0009	(-0.20)	0.0590	(7.21)
Seating capacity (SC level > status quo level) (positive) Hybrid	1.21	-0.0280	(-3.50)	0.1495	(9.25)
Seating capacity (SC level < status quo level) (negative) Hybrid	-1.31	-0.0133	(-1.65)	-0.0538	(-3.01)
Standard deviation of random parameters:					
Constant		-	-	0.2412	(3.00)
Gender (male=1)		-	-	0.3956	(5.62)
Number of acceptable alternatives per choice set		-	-	0.8035	(19.80)
Number of persons involved in recent purchase decision		-	-	0.0742	(1.93)
Australian manufactured vehicle (1,0)		-	-	0.0879	(1.19)
Personal income ($000s)		-	-	0.0128	(21.80)
Number years held driver's licence		-	-	0.0453	(21.70)
Full time employed (1,0)		-	-	0.6004	(10.10)
Part time employed (1,0)		-	-	0.3530	(3.44)
Climate change important issue		-	-	0.1002	(6.50)
People should be encouraged to use environmentally friend transport		-	-	0.4595	(21.90)
Govt. should implement carbon reduction policies		-	-	0.1640	(14.20)

Table 2 Summary of Model Results (cont.)

Drivers of high CO_2 cars should pay more		-	-	0.0977	(10.50)
Vehicle emissions charge is fair to all road users		-	-	0.5317	(24.30)
A vehicle emissions charge is effective way to reduce vehicle based CO2		-	-	0.5605	(23.60)
Household income ($000s)		-	-	0.0279	(24.80)
Number of children in household		-	-	0.0006	(3.27)
Number of adults in household		-	-	0.0020	(10.70)
Number of household members with a drivers licence		-	-	0.0015	(10.60)
Fuel cost (SC level < status quo level) (negative) Hybrid		-	-	0.0496	(0.83)
Regn cost (SC level > status quo level) (positive) Petrol		-	-	0.0014	(14.00)
Regn cost (SC level < status quo level) (negative) Petrol		-	-	0.0017	(9.06)
Regn cost (SC level < status quo level) (negative) Hybrid		-	-	0.0035	(12.80)
Fuel efficiency (SC level > status quo level) (positive) Diesel		-	-	0.0281	(1.90)
Fuel efficiency (SC level < status quo level) (negative) Diesel		-	-	0.2165	(16.80)
Seating capacity (SC level > status quo level) (positive) Hybrid		-	-	0.3314	(16.70)
Seating capacity (SC level < status quo level) (negative) Hybrid		-	-	0.0222	(1.62)
Variance function					
Number of acceptable alternatives per choice set	1.592	-0.2358	(-17.80)	-	-
Mean threshold parameters:					
μ_1		0.2552	(18.70)	-0.2986	(-6.43)
μ_2		0.5325	(27.50)	0.0433	(1.06)
μ_3		0.8786	(32.90)	0.6272	(16.80)
μ_4		1.3171	(36.30)	0.8215	(23.70)
μ_5		1.7354	(37.60)	1.0264	(27.50)
Standard deviation of random thresholds::					
μ_1		-	-	0.0915	(1.55)
μ_2		-	-	0.1668	(4.35)
μ_3		-	-	0.5419	(17.90)
μ_4		-	-	0.2346	(9.56)
μ_5		-	-	0.1299	(5.39)
Systematic sources of variation in random thresholds:					
Number of acceptable alternatives per choice set		-	-	0.0779	(11.20)
Heteroscedasticity in latent regression:					
Number of acceptable alternatives per choice set		-	-	-0.6820	(-35.70)
Latent heterogeneity in variance of random error (tau):		-	-	0.8944	(50.50)
Log-likelihood at zero		-10696.46			
Log-likelihood at convergence		-10315.75		-7930.23	
ρ^2		0.036		0.231	

Given the interest in the role of design dimensionality in influencing response choice certainty level, we begin by noting that the deviations of the attribute levels associated with the designed alternative from the reference or status quo levels for a number of the design attributes have a statistically significant influence on the choice certainty response, reinforcing the evidence in previous studies such as Brouwer et al. (2010) and Olsen et al. (2011). For example, the parameter estimate associated with fuel efficiency for a diesel vehicle when the

SC level is greater than the status quo level, is 0.0170 (*t*-ratio of 2.12), and when the SC level is less than the status quo level it is 0.0590 (*t*-ratio of 7.21). This suggests that, although the influence is directionally asymmetric, the greater the difference away from the status quo, the higher is the probability of greater surety about the choice response (resulting from multiplying the marginal utility of the attribute by the difference level), with the probability of surety being greater, for a given difference, when the SC level is less (i.e., more appealing) than the status quo. A possible interpretation of this finding is that the greater difference brings clarity of separation amongst the alternatives, increasing the certainty of the choice response. Furthermore, the positive sign in both directions suggests that there is a specific level of fuel efficiency that is preferred, that is offered by the status quo, and that higher or lower fuel efficiency implies something about the type of vehicle and hence some impact of response certainty.

Not all asymmetric attribute deviations from the status quo have a positive parameter estimate in one or both directions. Looking at seating, a positive (improvement) shift from the status quo yields an estimate of +0.1495, implying that the larger the positive shift the more certain is the respondent of the choice. On the other hand, the negative (deteriorated) shift from the SQ seating yields an estimate equal to -0.0538, implying less certainty as the attribute moves downwards from the SQ. A negative parameter suggests that the greater difference, which widens the gap between the design attribute level and the status quo level, regardless of whether positive or negative, lowers the probability of response certainty, which is also a plausible interpretation in terms of higher risk of moving to an alternative which is further away from the experienced alternative. Another interpretation of a negative parameter estimate is that a narrowing of the difference between the design attribute and status quo levels, while still reducing the probability of choice certainty, reduces this certainty by less than a wider range, on the argument that there is less risk in choosing a non-status quo alternative when it is closer to the status quo specification.

These deviation attributes are all random parameters, with statistically significant standard deviation parameters (except for the negative hybrid fuel cost), suggesting the presence of preference heterogeneity and hence variations in the probability of response certainty for a given deviation from the status quo level. We undertook a statistical test to establish if the asymmetry in parameter estimates associated with differencing around the status quo effects is statistically significant. This test can be undertaken for three pairs of estimates (namely, registration cost for petrol, fuel efficiency for diesel and seating capacity for hybrids). The evidence results in t-values of differences respectively of 3.28, 3.67 and 8.44, all significant at well above the 95 percent level of confidence.

The random threshold parameters are all statistically significant except for the mean threshold 2 and the standard deviation threshold 1. There is evidence of threshold heterogeneity across the sample, with the differences in the contribution of each threshold to the utility of each level of choice certainty being non-linear in recognition that the differences between the same absolute difference levels on the certainty scale translates into different utilities.

Furthermore, we have identified a systematic source of influence on these thresholds and the preference expression, defined as ‘the number of acceptable alternatives per choice set’ (as identified by an additional response to each choice set - see Figure 1). The highly significant positive parameter associated with the thresholds (0.0799 and *t*-ratio of 11.2) suggests that as the number of acceptable alternative increases, the threshold utility increases for each and every level of the choice certainty scale, which is added to the positive mean parameter effect of 0.3128 (and standard deviation parameter of 0.8035) in the preference expression. There is also a statistically significant mean parameter estimate in the preference expression (0.3128 with a *t*-ratio of 7.13), and a very significant standard deviation parameter estimate of 0.8035 (*t*-ratio of 19.8), which suggests that choice certainty increases when the number of acceptable alternatives increases. This is an interesting finding, since one may have thought that the fewer the number of acceptable alternatives, the easier it is to chose with certainty; however the greater variety of options (*within the limit of a maximum of three alternatives*) appears to give greater confidence in finding an alternative that is more acceptable. We suspect that the former argument would have greater currency when the number of alternatives become somewhat larger, but given that most choice experiments are limited to between two and four alternatives, the positive parameter estimate seems very plausible.

We also found that this same variable has a statistically significant and negative parameter estimate as a systematic source of influence on heteroscedasticity in the latent regression of random error component, $\sigma(i) = \exp[-0.6820 \times aacset + 0.8944 \times v(i)]$ {where aacset is the number of acceptable alternatives and *v(i)* is normally distributed, and 0.8944 is latent heterogeneity in variance of random error (τ in Table 2 and equation (10))}, indicating that as the number of acceptable alternatives per choice set increases, the residual unexplained utility decreases, which is a very plausible and pleasing finding. All of these findings integrate into equation (11), given equation (10), to impact on the probability of choice response certainty in a way that increases this probability as the number of acceptable alternatives increases.

There are a large number of socioeconomic characteristics that have a statistically significant discriminating role in the probability of a specific level of choice certainty. Beginning with gender, we see a significant parameter for the standard deviation estimate but not for the mean estimate, suggesting a large amount of respondent preference heterogeneity that is not captured by the mean. Other statistically significant positive effects are full time employment status, number of years that a driver's licence has been held; and negative effects are the number of children in the household and household income. The possible implication of this evidence is that the larger household size and wealth exerts a greater degree of uncertainty, which is in part offset by the greater certainty when the respondent is full time employed and has been driving for a longer period of time (the latter a possible proxy for experience in using a greater number of vehicle types). Among the socioeconomic attributes, some of the findings seem to further validate the GOL model, such as 'years of driving licence' having a significant positive effect in this model (which appears very intuitive). This evidence is reinforced in other studies cited in previously cited papers concerning the role of experience in increasing response certainty.

In addition to the socioeconomic influences, we investigated the role of attitudes to vehicle emissions. On a 7-point scale from strongly agree (7) to strongly disagree (1), we sought opinions on seven issues: "climate change is an important issue, vehicles are a main cause of climate change, people should be encouraged to use environmentally friendly transport, government should implement carbon reduction policies, drivers of high CO_2 cars should pay more, vehicle emissions charge is fair to all road users, and a vehicle emissions charge is an effective way to reduce vehicle based CO_2." The partial correlations between pairs of attitudinal questions are in the range 0.32 to 0.58. All but the second attitudinal issue (i.e., 'vehicles are a main cause of climate change') were statistically significant, with three attitudinal variables having negative mean estimates, and three having positive mean estimates.

Given that the certainty scale is an ordered scale of surety, the expected sign on the attitudinal variables is not intuitive. For example, a negative sign suggests, holding other effects constant, that a higher level of agreement in respect of the statement is associated with a lower level of utility associated with surety on the certainty scale. Although we are focussing on our preferred model, the GOL form, it is worth noting that there are sign changes in the attitudinal variables between the two models in Table 2. The change in the sign on the mean parameter estimate between the ordered and generalised ordered logit models appears to be due, in the main, to the account taken in the generalised model of the distribution of taste parameters which can change sign across the full distribution.

We find that respondents who tend to agree more on 'climate change as an important issue', 'that government should implement carbon reduction policies', and 'that drivers of high CO_2 cars should pay more', are more certain about their choice; in contrast respondents who tend to agree more that 'people should be encouraged to use environmentally friendly transport', 'that a vehicle emissions charge is fair to all road users', and 'that a vehicle emissions charge is an effective way to reduce vehicle based CO_2' are less certain about their choice. This reduced certainty might reflect greater ambiguity with the notion of an emissions charge in contrast to the perceived clarity of the climate change issue, carbon reduction policies and high emitting cars. However, despite the statistical significance of the parameter estimates, the influence of each attitude variable on surety is very small indeed when we feed in the attitudinal levels across the 7-point scale and assess changes in the probability of choice certainty.

5. Conclusions

This paper has investigated sources of systematic influence on the perceived certainty associated with the choice stated in a choice experiment. Although we might qualify the extent to which a certainty response on a surety scale is free of error, it nevertheless might be expected to offer some relevant information that is not on offer when we assume full certainty if nothing is known about the external validity of stated choice responses.

We have focused on an exploration of the role that the stated choice design itself might play in inducing variation in choice certainty. The three obvious candidate dimensions are the attribute levels associated with each alternative, how these relate to the levels of an experienced choice that is the pivot from which the design attribute levels are constructed (when such attributes are defined in an existing market for alternatives), and the acceptability of each alternative (itself associated with the attribute mix). There is clearly a connection between the perception of acceptable alternatives and the deviation of design attribute levels around the status quo levels that are associated with the experienced alternative.

The estimated model provides strong evidence that these design dimensions do influence the degree of choice certainty, and that the influence is spread throughout a number of elements of the utility expressions being represented in a generalized ordered logit choice model; notably through random parameters representing preference heterogeneity, through heteroscedasticity embedded in the random error, and via the random thresholds themselves conditioned on systematic sources of variation across the sampled respondents.

Given the statistically significant connection between SC design and choice certainty, what does this evidence mean for the future design of choice experiments, given an interest on increasing subjective choice response certainty? The clearest evidence is that choice experiments that are linked to a reference alternative provide a mechanism for at least assessing the extent to which design attribute level deviations condition the degree of choice certainty; however the fact that a narrower and a wider deviation both can be significant and plausible in increasing the probability of choice response certainty does not deliver guidelines on appropriate attribute levels and range.

We also recommend that accounting for the perceived acceptability of each alternative and hence the number of alternatives that are acceptable is essential information, which Hensher and Rose (*forthcoming*) have shown has a significant influence on the improvement in predictive power of choice models, and which clearly influences the certainty of choice response. Indeed we would go so far to suggest (given the evidence herein and in Hensher and Rose (*forthcoming*)) that conditioning a choice model on a knowledge of respondent perception of acceptable alternatives is something that is unambiguous in its impact of the predictive power of a choice model, as well as on the confidence we can associate with the improvement in predictive power in its link to increased choice certainty, as shown in the current paper. Hensher and Rose (*forthcoming*) incorporate this effect in the utility expression as a heteroscedastic conditioning on the form in (12, in summation notation).

$$U_{ji} = (1+\delta_j(AC_{ji} + \sum_{h=1}^{H} \gamma R_{hi})[\alpha_j + \sum_{k=1}^{K} \beta_{kj} X_{kji}] \qquad (12)$$

where AC_{jq} is a variable denoting whether an alternative is perceived to be acceptable or not by the q^{th} individual, R_{hq} is a dummy variable indicating whether the h^{th} attribute level is in a perceived attribute threshold rejection region or not for the q^{th} individual, and δ_j and γ_h are estimated parameters, α_j is an alternative-specific constant, and β_{kj} are the preference parameters associated with the k^{th} attribute (X) and j^{th} alternative. The inclusion of R_{hi} recognises that the role of attributes is fundamental to the perception of alternative acceptability.

Ongoing research should recognise the potential role that supplementary information on choice certainty might play in improving the external validity of probability outcomes associated with stated choice experiments. It would be especially encouraging if future research could undertake tests of choice certainty in the context of existing alternatives where there is an observed market choice from a set of alternatives that all currently exist, even if some attribute levels might be totally replicated in the stated choice and real world settings. A recommendation for future choice experiment designs flowing from this research and Hensher and Rose (*forthcoming*), is to select a choice set size so as to maximise the possibility that respondents will find acceptable alternatives (and be sure about choices).

Finally, we have included the attitudinal variables (or latent attributes) directly into the GoL mode; however an alternative approach that is gaining popularity is to jointly model choice and attitudes in what is referred to as a hybrid choice model (e.g., Ben-Akiva et al. 2002 and Bolduc et al. 2005), or in recent times as a choice model with latent variable scaling (see Hess and Hensher 2012).

Appendix

Table A1 Levels for Annual Emissions Surcharge ($)

Petrol		Fuel Efficiency (litres used per 100km)									
		6	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Level	**1**	0	0	0	0	0	0	0	0	0	0
	2	90	105	120	135	150	165	180	195	210	225
	3	180	210	240	270	300	330	360	390	420	450
	4	270	315	360	405	450	495	540	585	630	675
	5	360	420	480	540	600	660	720	780	840	900

Diesel		Fuel Efficiency (litres used per 100km)									
		6	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Level	**1**	0	0	0	0	0	0	0	0	0	0
	2	75	87.5	100	112.5	125	137.5	150	162.5	175	187.5
	3	150	175	200	225	250	275	300	325	350	375
	4	225	262.5	300	337.5	375	412.5	450	487.5	525	562.5
	5	300	350	400	450	500	550	600	650	700	750

Table A2 Levels for Variable Emissions Surcharge

Petrol		Fuel Efficiency (litres used per 100km)									
		6	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Level	**1**	0	0	0	0	0	0	0	0	0	0
	2	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15
	3	0.12	0.14	0.16	0.18	0.20	0.22	0.24	0.26	0.28	0.30
	4	0.18	0.21	0.24	0.27	0.30	0.33	0.36	0.39	0.42	0.45
	5	0.24	0.28	0.32	0.36	0.40	0.44	0.48	0.52	0.56	0.60

Diesel		Fuel Efficiency (litres used per 100km)									
		6	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Level	**1**	0	0	0	0	0	0	0	0	0	0
	2	0.05	0.06	0.07	0.08	0.09	0.09	0.10	0.11	0.12	0.13
	3	0.10	0.12	0.14	0.15	0.17	0.19	0.20	0.22	0.24	0.26
	4	0.15	0.18	0.2	0.23	0.26	0.28	0.31	0.33	0.36	0.38
	5	0.20	0.24	0.27	0.31	0.34	0.37	0.41	0.44	0.48	0.51

Source: Beck et al. (2011)

Acknowledgements

We acknowledge the significant contribution of Jun Zhang for his expert programming skills in developing the complex internet-based survey instrument in the ongoing automobile project and discussions with Joffre Swait, Waiyan Leong, Andrew Collins, Ricardo Scarpa and Benjamin McNair. This research was funded under an Australian Research Grant ARC DP0770618. We are indebted to two referees for very insightful comments.

References

Bech, M., T. Kjaer and J., Lauridsen, 2011. Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. *Health Economics*, 20(3), 273-286.

Beck, M. J., D. A., Hensher and J. M. Rose, 2009. *Report of a Pilot Survey for the Automobile Choice Project*. Institute of Transport and Logistics Studies, University of Sydney, May.

Beck, M. J., J. M., Rose and D. A., Hensher, 2011. Behavioural responses to vehicle emissions. *Transportation*, 38(3), 445-63.

Ben-Akiva, M., J., Walker, A., Bernardino, D., Gopinath, T., Morikawa and A., Polydoropoulou, 2002. Integration of choice and latent variable models. In Mahmassani, H. (Ed.), In *Perpetual motion: Travel behaviour research opportunities and application challenges*. Pergamon, Ch. 13, 431_470.

Bolduc, D., M., Ben-Akiva, J., Walker, A. and Michaud, 2005. Hybrid choice models with logit kernel: Applicability to large scale models. In Lee-Gosselin, M., Doherty, S. (Eds.), *Integrated Land-Use and Transportation Models: Behavioural Foundations.* Elsevier, Oxford, 275-302.

Brouwer, R., T., Dekker, J., Rolfe and J., Windle, 2010. Choice Certainty and Consistency in Repeated Choice Experiments. *Environmental and Resource Economics*, 46(1), 93-109.

Cantillo, V., B. Heydecker and J. de D., Ortuzar, 2006.A discrete choice model incorporating thresholds for perception in attribute values. *Transportation Research Part B*, 40(9), 807-825.

Champ, P. A. and R. C., Bishop, 2001. Donation payment mechanism and contingent valuation: An empirical study of hypothetical bias. *Environmental and Resource Economics*, 19(4), 383-402.

Chiou, L. and J. L., Walker, 2007. Masking Identification of Discrete Choice Models under Simulation Methods. *Journal of Econometrics*, 141(2), 683-703.

Fujii, S. and T. Gärling 2003. Application of attitude theory for improved predictive accuracy of stated preference methods in travel demand analysis. *Transportation Research Part A,* 37(4), 389-402.

Greene, W.H. and D. A., Hensher, 2010. *Modeling Ordered Choices: A Primer and Recent Developments*, Cambridge University Press, Cambridge.

Hensher, D. A. 2010. Hypothetical bias, choice experiments and willingness to pay. *Transportation Research Part B*, 44(6), 735-752.

Hensher, D. A. and A. T., Collins, 2011. Interrogation of responses to stated choice experiments: is there sense in what respondents tell us? A closer look at what respondents choose in stated choice experiments. *Journal of Choice Modelling,* 4(1), 62-89.

Hensher, D. A. and J. M. Rose, *forthcoming*. The Influence of Alternative Acceptability, Attribute Thresholds and Choice Response Certainty on Automobile Purchase Preferences, *Journal of Transport Economics and Policy*.

Hensher, D. A., and M. J., Beck and J. M., Rose, 2011. Accounting for preference and scale heterogeneity in establishing whether it matters who is interviewed to reveal household automobile purchase preferences. *Environment and Resource Economics*, 49(1), 1–22.

Hess, S. and D. A., Hensher, 2012. Making use of respondent reported processing information to understand attribute importance: a latent variable scaling approach. Paper presented at the 2012 *Transportation Research Board Annual Conference*, Washington DC.

Johannesson, M., G., Blomquist, K., Blumenschien, P., Johansson, B., Liljas and R., O'Connor, 1999. Calibrating hypothetical willingness to pay responses. *Journal of Risk and Uncertainty*, 8(1), 21-32.

Leung, W. and D. A., Hensher, 2011. Embedding multiple heuristics into choice models: an alternative approach. Presented at the *Second International Choice Modelling Conference*, Leeds, UK July 2011.

Li, C. and L., Mattsson, 1995. Discrete choice under preference uncertainty: an improved structural model for contingent valuation. *Journal of Environmental Economics and Management*, 28(2), 256-269.

Loomis, J. and E., Ekstrand, 1998. Alternative approaches for incorporating respondent uncertainty when estimating willingness to pay: the case of the Mexican spotted owl. *Ecological Economics*, 27(1), 29-41.

Lundhede T.H., S.B., Olsen, J.B., Jacobsen and J., Thorsen, 2008. Choice experiments and certainty in choice: a test of the influence of utility difference on self-reported certainty levels and evaluations of three recoding approaches to handle uncertain responses. Paper presented at the *16th Annual Conference of the European Association of Environmental and Resource Economists* (EAERE), June 25–28, Gothenburg

Lundhede, T., S., Olsen, J. Jacobsen, and B., Thorsen, 2009. Handling respondent uncertainty in Choice Experiments: Evaluating recoding approaches against explicit modelling of uncertainty. *Journal of Choice Modelling,* 2(2), 118-147.

Olsen, S., T. H., Lundhede, J. B., Jacobsen and B. J., Thorsen, 2011. Tough and Easy Choices: Testing the Influence of Utility Difference on Stated Certainty-in-Choice in Choice Experiments. *Environmental and Resource Economics*, 49(4), :491–510.

Rose, J.M. and M. C. J., Bliemer, 2008. Stated preference experimental design strategies. In Hensher, D. A. and K. J., Button (eds.) *Handbook of Transport Modelling*, Elsevier, Oxford, Ch8, 151-180.

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

Samnaliev, M., T. H., Stevens and T., More, 2006. A comparison of alternative certainty calibration techniques in contingent valuation. *Ecological* Economics, 57(3), 507-519.

Swait, J., 2001. A Non-compensatory choice model incorporating attribute cut-offs, *Transportation Research Part B*, 35(10), 903-928.