ITLS

WORKING PAPER

ITLS-WP-08-16

Ordered choices and heterogeneity in attribute processing

By

William H Greene* and
David A Hensher

*Department of Economics
Stern School of Business
New York University

July 2008

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES
The Australian Key Centre in
Transport and Logistics Management

The University of Sydney
Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-08-16

TITLE: **Ordered choices and heterogeneity in attribute processing**

ABSTRACT: A growing number of empirical studies involve the assessment of influences on a choice amongst ordered discrete alternatives. Ordered logit and probit models are well known, including extensions to accommodate random parameters and heteroscedasticity in unobserved variance. This paper extends the ordered choice random parameter model to permit random parameterization of thresholds and decomposition to establish observed sources of systematic variation in the threshold parameter distribution. We illustrate the empirical gains of this model in the context of an individual's choice amongst unlabelled attribute packages of alternative tolled and non-tolled routes for the commuting trip, and the role that each attribute plays, in the sense of being ignored or not. The ordering represents the number of attributes attended to from the full fixed set. The evidence suggests that there is significant heterogeneity associated with the thresholds that can be connected to systematic sources associated with the respondent (i.e., gender) and the choice experiment (i.e., aggregation treatment of components of travel time).

KEY WORDS: ***Ordered choice, heterogeneous thresholds, random parameters, stated choice designs, information processing, ignoring attributes***

AUTHORS: William H Greene and David A Hensher

CONTACT: Institute of Transport and Logistics Studies **(C37)**
The Australian Key Centre in Transport Management
The University of Sydney NSW 2006 Australia

Telephone: +61 9351 0071
Facsimile: +61 9351 0088
E-mail: itlsinfo@itls.usyd.edu.au
Internet: http://www.itls.usyd.edu.au

DATE: July 2008

Acknowledgements Research funded under the Australian Research Council Large Grants Scheme, Grant A00103962 and Discover Program grant DP0770618 and the Behavioral Choice Group in the Faculty of Economics and Business at the University of Sydney. Discussions with Chandra Bhat are appreciated as is the support of Ken Train.

1. Introduction

A growing number of empirical studies involve the assessment of influences on a choice amongst ordered discrete alternatives. Ordered logit and probit models are well known, including extensions to accommodate random parameters and heteroscedasticity in unobserved variance (see, e.g., Greene 2007). The ordered choice model allows non-linear effects of any variable on the probabilities associated with each ordered level (see for example, Eluru, Bhat and Hensher 2008). However the traditional ordered choice model is potentially limited, behaviorally, in that it holds the threshold values to be fixed. This can lead to inconsistent (*i.e.,* incorrect) estimates of the effects of variables. Extending the ordered choice model to account for threshold random heterogeneity as well as underlying systematic sources of explanation for unobserved heterogeneity is a logical extension in line with the growing interest in choice analysis in establishing additional candidate sources of observed and unobserved taste heterogeneity.

A substantive application herein is used to illustrate the behavioral gains from generalizing the ordered choice model to accommodate random thresholds. It is focused on the role that information processing strategies play in conditioning the way in which individuals assess the attributes associated with choice alternatives offered in a stated choice experiment (see Hensher et al. 2005a, Hensher 2006b, 2008). We investigate the role of attribute processing strategies (APS) in an individual's choice amongst unlabelled attribute packages of alternative tolled and non-tolled routes for the commuting trip. The ordering represents one very specific APS, namely the number of attributes attended to from the full set. Despite a growing number of studies focusing on these issues (see for example Cantillo *et al.* 2006, Hensher 2006, Swait 2001, Campbell et al. 2008), the entire domain of every attribute is treated as relevant to some degree and included in the utility expressions for every individual. While acknowledging the extensive study of nonlinearity in attribute specification which permits varying marginal (dis)utility over an attribute's range, including account for asymmetric preferences under conditions of gain and loss (see Hess at al. 2008), this is not the same as establishing ex ante the extent to which a specific attribute might be totally excluded from consideration for all manner of reasons, including the impost of the design of a choice experiment when stated choice data is being used.

Most psychological theories of choice assume a dual-phase model of the decision-making process (Houston *et al.*1989, Kahneman and Tversky, 1979, Thaler, 1999). The first phase relates to the editing of the problem. The second phase relates to the evaluation of the edited problem. The main function of the editing operations is "to organize and reformulate the options so as to simplify subsequent evaluation and choice" (Kahneman and Tversky, 1979, p. 274). The main function of the evaluation operations is to select the preferred alternative. Similarly, in other behavioral paradigms such as the 'Cancellation and Focus Model of Choice' (Houston et al., 1989; Houston and Sherman, 1995, Bonini *et al.* 2004), it is assumed that people cancel features shared by the alternatives (within bounds that allow for just noticeable difference), and focus evaluation on the remaining attributes.

Referencing is a critical activity in the construction of behavioral reality that captures many of the elements of stage one editing, including cancellation and focus. Referencing helps shape the perspectives through which an individual sees the world, focusing attention on key elements within, involving processes of *inclusion* and *exclusion* as well as *emphasis,* and hence operates by biasing the cognitive processes of information by individuals (Hallahan 1999). We are interested in two dimensions of referencing – attributes and choice[1]. Attribute referencing entails accentuation of attributes of alternatives, ignoring other attributes and hence biasing information processing in terms of focal attributes. Referencing suggests circumstances in SC studies where the alternatives on offer that are contrasts to the experienced alternative (for example, a recent trip) that offer less attractive attribute levels such as travel times, are more likely to result in higher willingness to pay compared to relatively more attractive attribute scenarios.

The establishment of attribute inclusion/exclusion in making choices in a stated choice (SC) context is often associated with design dimensionality and the so-called *complexity* of the SC experiment (Hensher 2006a). It is typically implied that designs with more items to evaluate are more complex than those with less items[2] (for example, Arentze *et al*., 2003, Swait and Adamowicz 2001a, 2001b), impose cognitive burden, and are consequently less reliable, in a behavioral sense, in revealing preference information. This is potentially misleading, since it suggests that complexity is an artefact of the *quantity* of information, in contrast to the *relevance* of information.

We need a way of identifying what information (that is, attributes) is actually processed in arriving at a choice outcome and which is ignored. We recognise however that the process is inherently stochastic from an analyst's perspective, since we will never be able, with total certainty, to rely on a set of exogenous data items to elicit how an attribute is processed by each individual. This necessitates treating the processing of attributes as endogenous with the choice outcome so that unobserved influences of processing can also be accommodated, at least randomly.

One way to establish potential sources of influence on attending to attributes is to investigate the link between the design of the choice context, the individual's background and the inclusion/exclusion of specific attributes. An ordered choice model is an appropriate econometric form within which to study the influences on the number of attributes ignored from the full set on offer; which can then be embedded in a joint process and outcome choice model (see Hensher 2008)

The paper is organised as follows. The next section sets out the econometric specification of the generalised ordered choice model, focusing on the derivation of the random threshold structure and its behavioral appeal. We then introduce the empirical context used to test this new model, focusing on the design of the stated choice experiment and associated questions used to define the choice setting and the process used by each respondent in establishing relevance of each attribute. The empirical analysis that follows presents the estimated the models – a traditional model and the extended ordered choice model, together with the associated marginal effects that are the basis of behavioral assessment. The paper concludes with some observations on the merits of the extended model form.

---

[1] Hallahan (1999) presents seven dimensions, of which attributes and choice are only two. The others are situations, actions, issues, responsibility and news.

[2] Complexity also includes attributes that are lowly correlated, in contrast to highly correlated, the latter supporting greater ease of assessment in that one attribute represents other attributes.

2. Generalizations of the ordered choice model to accommodate preference heterogeneity

2.1 The ordered probit model

The ordered probit model was proposed by Zavoina and McElvey (1975) for the analysis of categorical, nonquantitative choices, outcomes and responses. Familiar applications now include bond ratings, discrete opinion surveys such as those on political questions, obesity measures (Greene et al. 2008), preferences in consumption, and satisfaction and health status surveys such as those analyzed by Boes and Winkelmann (2004, 2007).

The model foundation is an underlying random utility or latent regression model,

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

in which the continuous latent utility, yi* is observed in discrete form through a censoring mechanism (equation 2).

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

The model contains the unknown marginal utilities, $\boldsymbol{\beta}$, as well as J+2 unknown threshold parameters, μ_j, all to be estimated using a sample of n observations, indexed by $i = 1,...,n$. The data consist of the covariates, $\mathbf{x}_i$ and the observed discrete outcome, $y_i = 0,1,...,J$. The assumption of the properties of the "disturbance," ε_i, completes the model specification. The conventional assumptions are that ε_i is a continuous disturbance with conventional cdf, $F(\varepsilon_i|\mathbf{x}_i) = F(\varepsilon_i)$ with support equal to the real line, and with density $f(\varepsilon_i) = F'(\varepsilon_i)$. The assumption of the distribution of ε_i includes independence from (or exogeneity of) $\mathbf{x}_i$.

By the laws of probability, the probabilities associated with the observed outcomes are given as equation (3).

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

Several normalizations 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 (as we will) that $\mathbf{x}_i$ contains a constant term, we will require $\mu_0 = 0$. With a constant term present, if this normalization is not imposed, then adding any nonzero 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.

Since the data contain no unconditional information on scaling of the underlying variable, if y_i* is scaled by any positive value, then scaling the unknown μ_j and $\boldsymbol{\beta}$ by the same value preserves the observed outcomes – a free unconditional variance parameter, $\text{Var}[\varepsilon_i] = \sigma_\varepsilon^2$, is not identified without further restriction. We thus impose the identifying restriction σ_ε = a known constant, $\bar{\sigma}$. The usual approach to this normalization is to assume that $\text{Var}[\varepsilon_i|\mathbf{x}_i] = 1$ in the probit model and $\pi^2/3$ in the logit model – in both cases to eliminate the free structural scaling parameter. The standard treatments in the received literature complete the ordered choice model by assuming either a standard normal distribution for ε_i, producing the ordered probit model or a standardized logistic distribution (mean zero, variance $\pi^2/3$), which produces the ordered logit model. Applications appear to be well divided between the two. A compelling case for a particular distribution remains to be put forth.

With the full set of normalizations in place, the likelihood function for estimation of the model parameters is based on the implied probabilities given in equation (4).

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

Estimation of the parameters is a straightforward problem in maximum likelihood estimation (see, e.g., Greene 2008 and Pratt 1981). Interpretation of the model parameters is, however, much less so (see, e.g., Daykin and Moffitt 2002). There is no natural conditional mean function, so in order to attach behavioral meaning to the parameters, one typically refers to the probabilities themselves. The partial effects in the ordered choice model are:

$$\frac{\partial \text{Prob}[y = j \mid \mathbf{x}_i]}{\partial \mathbf{x}_i} = \left[f(\mu_{j-1} - \boldsymbol{\beta}'\mathbf{x}_i) - f(\mu_j - \boldsymbol{\beta}'\mathbf{x}_i) \right]\boldsymbol{\beta}. \tag{5}$$

The result shows that neither the sign nor the magnitude of a coefficient is informative about the corresponding behavioral characteristic in the model, so the direct interpretation of the coefficients (or their “significance”) is fundamentally ambiguous. A counterpart result for a dummy variable in the model would be obtained by using a difference of probabilities, rather than a derivative (Boes and Winkelmann 2007 and Greene 2008, Chapter E22). One might also be interested in cumulative values of the partial effects, such as shown in equation (6) (see, e.g., Brewer et al. 2006). The last term in this set is zero by construction.

$$\frac{\partial \text{Prob}[y \le j \mid \mathbf{x}_i]}{\partial \mathbf{x}_i} = \left(\sum_{m=0}^{j} \left[f(\mu_{m-1} - \boldsymbol{\beta}'\mathbf{x}_i) - f(\mu_m - \boldsymbol{\beta}'\mathbf{x}_i) \right] \right)\boldsymbol{\beta}. \tag{6}$$

2.2 A generalized ordered choice model

A number of authors, beginning with Terza (1985), have questioned some of the less flexible aspects of the model specification. Recent analyses, e.g., Long (1993), Long and Frees (2005) and Williams (2006), have proposed a “generalized ordered choice model." The partial effects shown above vary with the data and the parameters. Since the probabilities must sum to one, the partial effects for each variable must sum to zero across the probabilities. It can also be shown that for the probit and logit models, this set of partial derivatives will change sign exactly once in the sequence from 0 to J, a property that Boes and Winkelmann (2007) label the "single crossing" characteristic. Boes and Winkelmann (2007) also note that for any two continuous covariates, x_{ik} and x_{il}

$$\frac{\partial \text{Prob}[y_i = j \mid \mathbf{x}_i]/\partial x_{i,k}}{\partial \text{Prob}[y_i = j \mid \mathbf{x}_i]/\partial x_{i,l}} = \frac{\beta_k}{\beta_l} \tag{7}$$

This result in (7) is independent of the outcomes. The ordered choice models above have the property in equation (8); that is, the partial effects are each a multiple of the same $\boldsymbol{\beta}$.

$$\partial \text{Prob}[y_i > j \mid \mathbf{x}_i]/\partial \mathbf{x}_i = K_j \boldsymbol{\beta} \tag{8}$$

This is a feature of the model that has been labeled the "parallel regressions" assumption. A standard test of this null hypothesis, due to Brant (1990), is used to detect the condition. The Brant test frequently rejects the null hypothesis of a common slope vector in these discrete choice models. Since the discrete choices are not independent; indeed, the entire model describes a single choice (e.g., if y > 2, then y > 1), the test is not about the choice mechanism per se, but about functional form.

An extended form of the ordered choice model that has attracted much (perhaps most) of the recent attention, is the "Generalized Ordered Logit" (or Probit) model e.g., by Williams (2006). This model is defined in equation (9).

$$\text{Prob}[y_i = j \mid \mathbf{x}_i] = \text{Prob}[\varepsilon_i < \mu_j - \boldsymbol{\beta}_j'\mathbf{x}_i] - \text{Prob}[\mu_{j-1} - \boldsymbol{\beta}_{j-1}'\mathbf{x}_i], j = 0,1,..., J \tag{9}$$

where $\boldsymbol{\beta}_{-1} = 0$ (see e.g., Williams 2006, Long 1997, Long and Frees 2006). The extension provides for a separate vector of marginal utilities for each outcome.

The generalization of the model suggested above deals with both problems (single crossing and parallel regressions), but it creates new ones. The heterogeneity in the parameter vector is an artifact of the coding of the dependent variable, not a manifestation of underlying heterogeneity in the dependent variable induced by behavioral differences. It is unclear what it means for the marginal utility parameters to be structured in this way. Consider, for example, that there is no underlying structure that could be written down in such a way as to provide a means of simulating the data generating mechanism. By implication, $y_i^* = \boldsymbol{\beta}_j'\mathbf{x}_i + \varepsilon_i$ if $y_i = j$. That is, the model structure is endogenous – one could not simulate a value of y_i from the data generating mechanism without knowing in advance the value being simulated. There is no reduced form. The more difficult problem of this generalization is that the probabilities in this model need not be positive, and there is no parametric restriction (other than the restrictive model version we started with) that could achieve this. The probability model is internally inconsistent. The restrictions would have to be functions of the data. The problem is noted by Williams (2006), but dismissed as a minor issue. Boes and

Winkelmann (2007) suggest that the problem could be handled through a “nonlinear specification.” Essentially, this generalized choice model does not treat the outcome as a single choice, even though that is what it is.

To put a more positive view, we might interpret this as a semi-parametric approach to modeling what is underlying heterogeneity. However, it is not clear why this heterogeneity should be manifest in parameter variation across the outcomes instead of across the individuals in the sample. One would assume that the failure of the Brant test to support the model with parameter homogeneity is, indeed, signalling some failure of the model. A shortcoming of the functional form as listed above (compared to a different internally consistent specification) is certainly a possibility. We hypothesise that it might also be picking up unobserved heterogeneity.

2.3 Modeling observed and unobserved heterogeneity

Since Terza (1985), with the exception of Pudney and Shields (2000), most of the “generalizations” suggested for the ordered choice models have been about functional form – the single crossing feature and the parallel regressions (see, also, Greene 2008). Our interest in this paper is, rather, in a specification that accommodates both observed and unobserved heterogeneity across individuals. We suggest that the basic model structure, when fully specified, provides for sufficient nonlinearity to capture the important features of choice behavior. The generalization that interests us herein will incorporate both observed and unobserved heterogeneity in the model itself.

The basic model assumes that the thresholds μ_j are the same for every individual in the sample. Terza (1985), Pudney and Shields (2000), Boes and Winkelmann (2007) and Greene et al. (2008), all present cases that suggest individual variation in the set of thresholds is a degree of heterogeneity that is likely to be present in the data, but is not accommodated in the model. Pudney and Shields discuss a clear example in the context of job promotion, in which the steps on the promotion ladder for nurses are somewhat individual specific.

Greene (2002, 2008) argues that the fixed parameter version of the ordered choice model, and more generally, many microeconometric specifications, do not adequately account for the underlying, unobserved heterogeneity likely to be present in observed data. Further extensions of the ordered choice model presented in Greene (2008) include full random parameters treatments and discrete approximations under the form of latent class, or finite mixture models. These two specific extensions are also listed by Boes and Winkelmann (2004, 2007) as candidates for extending the model. They also describe a common effects model for panel data.

The model that assumes homogeneity of the preference parameters, $\boldsymbol{\beta}$, across individuals, also assumes homogeneity in the scaling of the random term, ε_i. That is, the homoscedasticity assumption, $\text{Var}[\varepsilon_i|\mathbf{x}_i] = 1$ is restrictive in the same way that the homogeneity assumption is. Heteroscedasticity in terms of observables in the ordered choice model is proposed in Greene (1997) and reappears as a theme in Williams (2006).

The model proposed here generalizes the ordered choice model in the directions of accommodating heterogeneity, rather than in the direction of adding nonlinearities to the underlying functional form. The earliest extensions of the ordered choice model focused on the threshold parameters. Terza’s (1985) extension suggested

$$\mu_{ij} = \mu_j + \boldsymbol{\delta}'\mathbf{z}_i. \tag{10}$$

The analysis of this model continued with Pudney and Shields's (2000) "Generalized Ordered Probit Model," whose motivation, like Terza's was to accommodate *observable* individual heterogeneity in the threshold parameters as well as in the mean of the regression. We (and Pudney and Shields) note an obvious problem of identification in this specification. Consider the generic probability with this extension,

$$\text{Prob}[y_i \leq j \mid \mathbf{x}_i,\mathbf{z}_i] = F(\mu_j + \boldsymbol{\delta}'\mathbf{z}_i - \boldsymbol{\beta}'\mathbf{x}_i) = F[\mu_j + (\boldsymbol{\delta}^{*\prime}\mathbf{z}_i + \boldsymbol{\beta}'\mathbf{x}_i)],\ \boldsymbol{\delta}^* = -\boldsymbol{\delta}. \tag{11}$$

It is less than obvious whether the variables $\mathbf{z}_i$ are actually in the threshold or in the mean of the regression. Either interpretation is consistent with the model. Pudney and Shields argue that the distinction is of no substantive consequence for their analysis.

Formal modeling of heterogeneity in the parameters as representing a feature of the underlying data, also appears in Greene (2002) (version 8.0) and Boes and Winkelmann (2004), both of whom suggest a random parameters (RP) approach to the model. In Boes and Winkelmann, it is noted that the nature of an RP specification induces heteroscedasticity, and could be modeled as such. The model would appear as follows:

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \mathbf{u}_i \tag{12}$$

where $\mathbf{u}_i \sim N[\mathbf{0},\boldsymbol{\Omega}]$. Inserting this in the base case model and simplifying, we obtain equation (13).

$$\text{Prob}[y_i \leq j \mid \mathbf{x}_i] = \text{Prob}[\varepsilon_i + \mathbf{u}_i'\mathbf{x}_i \leq \mu_j - \boldsymbol{\beta}'\mathbf{x}_i] = F\left(\frac{\mu_j - \boldsymbol{\beta}'\mathbf{x}_i}{\sqrt{1+\mathbf{x}_i'\boldsymbol{\Omega}\mathbf{x}_i}}\right), \tag{13}$$

Equation (13) could be estimated by ordinary means, albeit with a new source of nonlinearity – the elements of $\boldsymbol{\Omega}$ must now be estimated as well[3]. Boes and Winkelmann (2004, 2007) did not pursue this approach. Greene (2002) analyzes essentially the same model, but proposes to estimate the parameters by maximum simulated likelihood.

Curiously, none of the studies listed above focus on the issue of scaling, although Williams (2006), citing Allison (1999) does mention it. A heteroscedastic ordered probit model with the functional form in (14) appears at length in Greene (1997), and is discussed in some detail in Williams (2006).

$$\text{Var}[\varepsilon_i|\mathbf{h}_i] = \exp(\boldsymbol{\gamma}'\mathbf{h}_i) \tag{14}$$

[3] The authors' suggestion that this could be handled semiparametrically without specifying a distribution for u_i is incorrect, because the resulting heteroscedastic probability written above only preserves the standard normal form assumed if u_i is normally distributed as well as ε_i

In microeconomic data, scaling of the underlying preferences is as important a source of heterogeneity as displacement of the mean, perhaps even more so. But, it has received considerably less attention than heterogeneity in location.

In what follows, we will propose a formulation of the ordered choice model that treats heterogeneity in a unified, internally consistent fashion. The model contains three points at which individual heterogeneity can substantively appear: in the random utility model (the marginal utilities), in the threshold parameters, and in the scaling (variance) of the random components. As argued above, this form of treatment seems more likely to capture the salient features of the data generating mechanism than the received "generalized ordered logit model," which is more narrowly focused on functional form.

2.4 Random thresholds and heterogeneity in the ordered choice model

We depart from the base case,

$$\text{Prob}[y_i = j \mid \mathbf{x}_i] = F(\mu_j - \boldsymbol{\beta}'\mathbf{x}_i) - F(\mu_{j-1} - \boldsymbol{\beta}'\mathbf{x}_i) > 0, j = 0,1,...,J. \tag{15}$$

The intrinsic heterogeneity in utility functions across individuals is captured by writing

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\Delta}\mathbf{z}_i + \boldsymbol{\Gamma}\mathbf{v}_i \tag{16}$$

where $\boldsymbol{\Gamma}$ is a lower triangular matrix and $\mathbf{v}_i \sim \text{N}[\mathbf{0},\mathbf{I}]$. $\boldsymbol{\beta}_i$ is normally distributed across individuals with conditional mean (equation 17):

$$E[\boldsymbol{\beta}_i|\mathbf{x}_i,\mathbf{z}_i] = \boldsymbol{\beta} + \boldsymbol{\Delta}\mathbf{z}_i \tag{17}$$

and conditional variance (equation 18):

$$\text{Var}[\boldsymbol{\beta}_i|\mathbf{x}_i,\mathbf{z}_i] = \boldsymbol{\Gamma}\mathbf{I}\boldsymbol{\Gamma}' = \boldsymbol{\Omega}. \tag{18}$$

This is a random parameters formulation that appears elsewhere, e.g., Greene (2002, 2005). The random effects model is a special case in which only the constant is random. The Mundlak (1978) and Chamberlain (1980) approach to modeling fixed effects is also accommodated by letting $\mathbf{z}_\text{i} = \overline{\mathbf{x}}_i$ in the equation for the overall constant term.

The thresholds are modeled randomly and nonlinearly as

$$\mu_{ij} = \mu_{i,j-1} + \exp(\alpha_j + \boldsymbol{\delta}'\mathbf{r}_i + \sigma_j w_{ij}), w_{ij} \sim \text{N}[0,1] \tag{19a}$$

with normalizations and restrictions $\mu_{-1} = -\infty$, $\mu_0 = 0$, $\mu_J = +\infty$. For the remaining thresholds, we have (19).

$$\mu_1 = \exp(\alpha_1 + \boldsymbol{\delta}'\mathbf{r}_i + \sigma_1 w_{j1})$$

$$= \exp(\boldsymbol{\delta}'\mathbf{r}_i)\exp(\alpha_1 + \sigma_1 w_{j1}) \qquad (19b)$$

$$\mu_2 = \exp(\boldsymbol{\delta}'\mathbf{r}_i)\,[\exp(\alpha_1 + \sigma_1 w_{j1}) + \exp(\alpha_2 + \sigma_2 w_{j2})],$$

$$\mu_j = \exp(\boldsymbol{\delta}'\mathbf{r}_i)\left(\Sigma_{m=1}^{j}\exp(\alpha_m + \sigma_m w_{im})\right),\ j = 1,\ldots,J\text{-}1$$

$$\mu_J = +\infty.$$

This formulation preserves the ordering of the thresholds and incorporates the necessary normalizations. 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 ($\mathbf{r}_i$) and unobservable (w_{ij}) heterogeneity. The model is fully consistent, in that the probabilities are all positive and sum to one by construction. If $\boldsymbol{\delta} = \mathbf{0}$ and $\sigma_j = 0$, then the original model is returned, with $\mu_1 = \exp(\alpha_1)$, $\mu_2 = \mu_1 + \exp(\alpha_2)$ and so on.

Finally, the disturbance variance is allowed to be heteroscedastic, now specified randomly as well as deterministically. Thus,

$$\text{Var}[\varepsilon_i|\mathbf{h}_i,e_i] = \sigma_i^2 = \exp(\boldsymbol{\gamma}'\mathbf{h}_i + \tau e_i) \qquad (20)$$

where $e_i \sim N[0,1]$. Let $\mathbf{v}_i = (v_{i1},\ldots,v_{iK})'$ and $\mathbf{w}_i = (w_{i1},\ldots,w_{i,J\text{-}1})'$. Combining terms, the conditional probability of outcome j is

$$\text{Prob}[y_i = j \mid \mathbf{x}_i,\mathbf{z}_i,\mathbf{h}_i,\mathbf{r}_i,\mathbf{v}_i,\mathbf{w}_i,e_i] = F\left[\frac{\mu_{ij} - \boldsymbol{\beta}_i'\mathbf{x}_i}{\sqrt{\exp(\boldsymbol{\gamma}'\mathbf{h}_i + \tau e_i)}}\right] - F\left[\frac{\mu_{i,j-1} - \boldsymbol{\beta}_i'\mathbf{x}_i}{\sqrt{\exp(\boldsymbol{\gamma}'\mathbf{h}_i + \tau e_i)}}\right] \qquad (21)$$

The term that enters the log likelihood function is unconditioned on the unobservables. Thus,

$$\text{Prob}[y_i = j \mid \mathbf{x}_i,\mathbf{z}_i,\mathbf{h}_i,\mathbf{r}_i] =$$

$$\int_{\mathbf{v}_i,\mathbf{w}_i,e_i}\left(F\left[\frac{\mu_{ij} - \boldsymbol{\beta}_i'\mathbf{x}_i}{\sqrt{\exp(\boldsymbol{\gamma}'\mathbf{h}_i + \tau e_i)}}\right] - F\left[\frac{\mu_{i,j-1} - \boldsymbol{\beta}_i'\mathbf{x}_i}{\sqrt{\exp(\boldsymbol{\gamma}'\mathbf{h}_i + \tau e_i)}}\right]\right) f(\mathbf{v}_i,\mathbf{w}_i,e_i)\,d\mathbf{v}_i\,d\mathbf{w}_i\,de_i. \qquad (22)$$

The model is estimated by maximum simulated likelihood. The simulated log likelihood function is given in (23).

$$\log L_S(\boldsymbol{\beta},\boldsymbol{\Delta},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\gamma},\boldsymbol{\Gamma},\boldsymbol{\sigma},\tau) =$$

$$\sum_{i=1}^{n}\log\frac{1}{M}\sum_{m=1}^{M}\left(F\left[\frac{\mu_{ij,m} - \boldsymbol{\beta}_{i,m}'\mathbf{x}_i}{\sqrt{\exp(\boldsymbol{\gamma}'\mathbf{h}_i + \tau e_{i,m})}}\right] - F\left[\frac{\mu_{i,j-1,m} - \boldsymbol{\beta}_{i,m}'\mathbf{x}_i}{\sqrt{\exp(\boldsymbol{\gamma}'\mathbf{h}_i + \tau e_{i,m})}}\right]\right) \qquad (23)$$

$\mathbf{v}_{i,m}$, $\mathbf{w}_{i,m}$, $e_{i,m}$ are a set of M multivariate random draws for the simulation[4]. This is the model in its full generality. Whether a particular data set will be rich enough to support this much parameterization, particularly the elements of the covariances of the unobservables in $\boldsymbol{\Gamma}$, is an empirical question that will depend on the application.

The model contains four 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, $\boldsymbol{\beta}_i$, in the utility function, $\mathbf{x}_i$ and in the variance, σ_i^2. For convenience in the derivation below, let a vector $\mathbf{a}_i$ denote the union of $(\mathbf{x}_i,\mathbf{r}_i,\mathbf{z}_i,\mathbf{h}_i)$. This allows for cases in which variables appear at more than one place in the model. The partial effect of a change in an element of $\mathbf{a}_i$ on the probability will depend on where it appears in the specification. For cases in which a variable appears in more than one location, the partial effect will be the sum of the two, three or four terms. To avoid a cumbersome reparameterization of the model to place zeros in the appropriate places in the various parameter vectors and matrix, we assume at this point that $\mathbf{a}_i$ appears in full throughout the model; that is, as if $\mathbf{a}_i = \mathbf{x}_i = \mathbf{r}_i = \mathbf{z}_i = \mathbf{h}_i$. Thus, we write the probability of interest as (24).

$$\text{Prob}(y_i = j \mid \mathbf{a}_i) =$$
$$\int_{\mathbf{v}_i,\mathbf{w}_i,e_i} \left(F\left[\frac{\mu_{ij} - (\beta + \Delta\mathbf{a}_i + \boldsymbol{\Gamma}\mathbf{v}_i)'\mathbf{a}_i}{\sqrt{\exp(\gamma'\mathbf{a}_i + \tau e_i)}}\right] - F\left[\frac{\mu_{i,j-1} - (\beta + \Delta\mathbf{a}_i + \boldsymbol{\Gamma}\mathbf{v}_i)'\mathbf{a}_i}{\sqrt{\exp(\gamma'\mathbf{a}_i + \tau e_i)}}\right]\right) f(\mathbf{v}_i,\mathbf{w}_i,e_i)d\mathbf{v}_i d\mathbf{w}_i de_i. \tag{24}$$

μ_{ij} is defined in (19). Then, the set of partial effects is given as (25).

$$\frac{\partial \text{Prob}(y_i = j \mid \mathbf{a}_i)}{\partial \mathbf{a}_i} =$$
$$\int_{\mathbf{v}_i,\mathbf{w}_i,e_i} \left(f\left[\frac{\mu_{ij} - \beta_i'\mathbf{a}_i}{\sqrt{\exp(\gamma'\mathbf{a}_i + \tau e_i)}}\right] \frac{1}{\sqrt{\exp(\gamma'\mathbf{a}_i + \tau e_i)}} \left(\beta_i + 2\Delta\mathbf{a}_i - \tfrac{1}{2}\gamma + \mu_{ij}\delta\right)\right) f(\mathbf{v}_i,\mathbf{w}_i,e_i)d\mathbf{v}_i d\mathbf{w}_i de_i$$
$$-\int_{\mathbf{v}_i,\mathbf{w}_i,e_i} \left(f\left[\frac{\mu_{i,j-1} - \beta_i'\mathbf{a}_i}{\sqrt{\exp(\gamma'\mathbf{a}_i + \tau e_i)}}\right] \frac{1}{\sqrt{\exp(\gamma'\mathbf{a}_i + \tau e_i)}} \left(\beta_i + 2\Delta\mathbf{a}_i - \tfrac{1}{2}\gamma + \mu_{i,j-1}\delta\right)\right) f(\mathbf{v}_i,\mathbf{w}_i,e_i)d\mathbf{v}_i d\mathbf{w}_i de_i \tag{25}$$

The sum of four terms in the middle of the expressions shows the four parts of a compound partial effect; in turn, these are the components of the change (a) due directly to change in x_i, (b) indirectly due to change in the variables that influence β_i, (c) due to change in the variables in the variance and (d) due to changes in the threshold parameters, respectively.

Like the log likelihood function, the partial effects must be computed by simulation. If a variable appears only in $\mathbf{x}_i$, then this formulation retains both the "parallel regressions" and "single crossing" features of the original model. Nonetheless, the effects are highly nonlinear in any event. However, if a variable appears anywhere else in the specification, then neither of these properties will remain.

[4] We use Halton sequences rather than pseudo-random numbers. See Train (2003) for discussion.

3. Empirical application

The context of the application, using stated choice data from a larger study reported in Hensher (2006a,b), is an individual's choice amongst unlabelled attribute packages of alternative tolled and non-tolled routes for the car commuting trip in Sydney (Australia) in 2002. In this paper we are interested in one feature of the way in which individual's process attribute information, namely attribute inclusion or exclusion. The dependent variable in the ordered choice model is the number of ignored attributes, or the number of attributes attended to from the full fixed set associated with each alternative package of route attributes. The utility function is defined over the attribute information processed by each individual, with candidate influences on the each individual's decision heuristic including the dimensions of the choice experiment (e.g., number of alternatives, range of attributes), the framing of the design attribute levels relative to a reference alternative (see below), an individuals socioeconomic characteristics, and attribute accumulation where attributes are in common units (see also Hensher 2006b).

The alternative attribute packages offered to individuals to evaluate are pivoted around the car commuting experiences of sampled respondents. The use of a respondent's experience, embodied in a reference alternative, to derive the attribute levels of the experiment is supported by a number of theories in behavioural and cognitive psychology, and economics, such as prospect theory, case-based decision theory and minimum-regret theory (see Starmer 2000, Hensher 2006b).

Reference alternatives in SC experiments[5] act to frame the decision context of the choice task within some existing memory schema of the individual respondents and hence make preference-revelation more meaningful at the level of the individual. Theoretically, the role of reference alternatives in SC tasks is well supported within the literature. For example, prospect theory (Kahneman and Tversky 1979), which argues that individuals use decision heuristics when making choices, promotes the idea that the very specific context in which a decision is made by each individual is an important determinant of the selection of choice-heuristic, supporting the use of reference alternatives in SC tasks. Framing effects, of which reference dependence is a popular interpretation, provides context support in trading off the desire to make a good choice against the cognitive effort involved in processing the additional information provided in a SC task (Hensher 2006). Starmer (2000, p 353) in particular argues strongly for the use of reference alternatives (e.g., a current trip) in decision theory.

16 stated choice sub-designs have been embedded in one overall design. Each commuter evaluated one randomly assigned sub-design; however, across the full set of stated choice experiments, the designs differed in terms of the number, range and levels of attributes, the number of alternatives and the number of choice sets. The combination of these dimensions of each design is often seen as the source of design 'complexity', and it is within this setting that we have varied the number of attributes that each respondent is asked to evaluate, and through supplementary questions, established which attributes were 'ignored' in the evaluation and selection of an alternative.

Previous studies were used to identify candidate design dimensions. The five design dimensions are shown in Table 1.

[5] Hensher (2004), Train and Wilson (2008), and Rose et al. (2008) provide details of the design of pivot-based experiments.

Table 1: Dimensionality of the design plan

Choice set size	Number of alternatives	Number of attributes	Number of attribute levels	Range of attribute levels
6	2	3	2	Narrower than base
9	3	4	3	Base
12	4	5	4	Wider than base
15	----	6	----	----

Six attributes were selected for each alternative, based on previous evidence (see Hensher 2001), to characterise the options: free-flow time, slowed down time, stop/start time, variability of trip time, toll cost and running costs. Hensher (2006) explored how varying the number of attributes affects information processing, grouping attributes according to four patterns, noting that aggregated attributes are combinations of existing attributes[6].

In the current study we focus on one of the four patterns, that associated with five attributes; i.e., free flow time, slowed down time, stop/start time, trip time variability, total costs. We have selected a generic design (i.e., unlabeled alternatives) to avoid confounding the effect of the number of alternatives with the labeling (e.g., car, train). The sub-design dimensions are shown in Table 2 with eth attribute ranges in Table 3.

Table 2: The sub-designs of the overall design for five attributes

Choice set of size	Number of alternatives	Number of attributes	Number of levels of attributes	Range of attribute levels
15	2	5	2	Wider than base
9	2	5	4	Base
6	3	5	4	Narrower than base
12	4	5	3	Narrower than base

Note: Column 1 refers to the number of choice sets. The four rows represent the set of designs (see Appendix A).

Table 3: The attribute profiles for the design

(units = %)	Base range			Wider range			Narrower range		
Levels:	2	3	4	2	3	4	2	3	4
Free flow time	± 20	-20, 0, +20	-20,-10,+10,+20	-20, +40	-20,+10,+40	-20, 0,+20,+40	± 5	-5, 0,+5	-5, -2.5, +2.5, +5
Slow down time	± 40	-40, 0, +40	-40,-20,+20,+40	-30, +60	-30,+15,+60	-30, 0,+30,+60	± 20	-20, 0, +20	-20, -2.5, +2.5, +20
Stop/start time	± 40	-40, 0, +40	-40,-20,+20,+40	-30, +60	-30,+15,+60	-30, 0,+30,+60	± 20	-20, 0, +20	-20, -2.5, +2.5, +20
Uncertainty of travel time	± 40	-40, 0, +40	-40,-20,+20,+40	-30, +60	-30,+15,+60	-30, 0,+30,+60	± 20	-20, 0, +20	-20, -2.5, +2.5, +20
Total costs	± 20	-20, 0, +20	-20,-10,+10,+20	-20, +40	-20,+10,+40	-20, 0,+20,+40	± 5	-5, 0,+5	-5, -2.5, +2.5, +5

As a generic design, the added alternatives are exactly the same. That is, for two design alternatives, we should not expect to find the parameter for an attribute (e.g., 'free flow travel time') to be different for the set of non-reference alternatives. Therefore we do not need the attribute 'free flow time one' to be orthogonal to the attribute 'free flow time two' etc up to 'free flow time *J*-1'. We need to ensure that the attribute 'free flow time' representing all non-reference alternatives is perfectly[7] orthogonal to the other attributes (such as slow down time, etc.). The designs are computer-generated. A

[6] This is an important point because we did not want the analysis to be confounded by extra attribute dimensions.

[7] *Approximately* orthogonal is also acceptable given that some designs cannot guarantee complete orthogonality without loss of structure in terms of cognitive efficiency (in contrast to statistical efficiency).

preferred choice experiment design is one that maximizes the determinant of the covariance matrix, which is itself a function of the estimated parameters. Knowledge of the parameters or at least some priors (such as signs) for each attribute, from past studies, provides a useful input. We found that in so doing, the search eliminates dominant alternatives. The method used finds the D-optimality plan very quickly (see Rose and Bliemer 2007).

The *actual* levels of the attributes shown to respondents are calculated relative to those of the experienced reference alternative – a recent car commuter trip. The levels applied to the choice task differ depending on the range of attribute levels and the number of levels for each attribute. The design dimensions are translated into SC screens, illustrated in Figure 1. The range of the attribute levels vary *across* designs. Each sampled commuter is given a varying number of choice sets (or scenarios), but the number of attributes and alternatives remain fixed. Elicitation questions associated with attribute inclusion and exclusion shown in Figure 2.

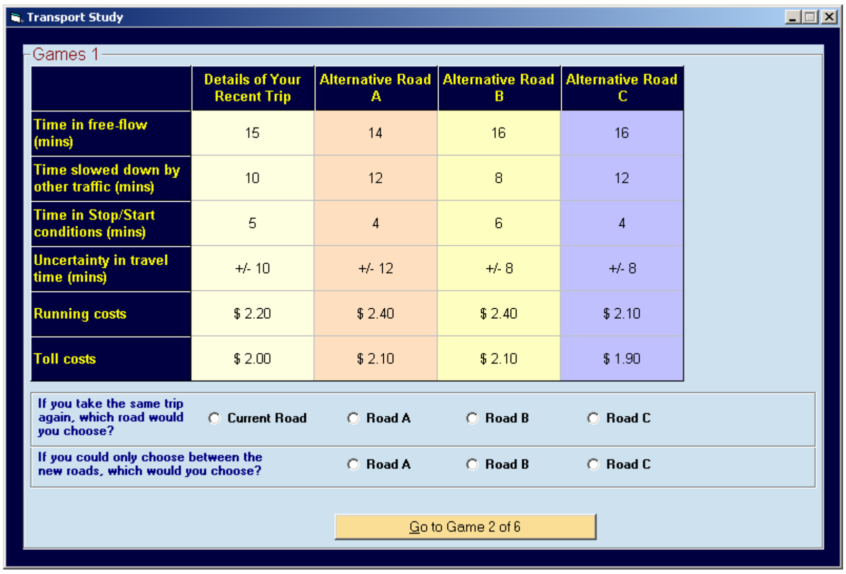

Figure 1: An example of a stated choice screen

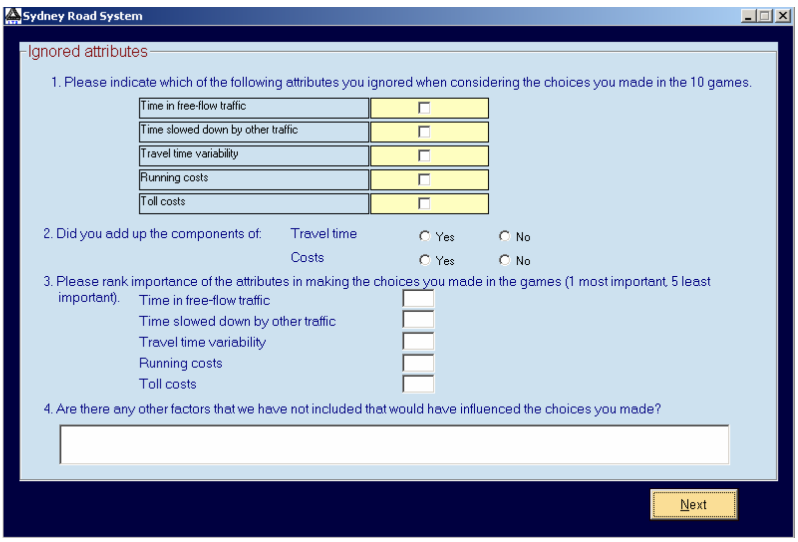

Figure 2: CAPI questions on attribute relevance

4. Empirical analysis

Computer-aided personal interview (CAPI) surveys were completed in the Sydney metropolitan area in 2002. A stratified random sample was applied, based on the residential location of the household. Screening questions established eligibility in respect of commuting by car. Further details are given in Hensher (2006a). Final models are given in Table 4, with the respective marginal effects in Table 5.

A direct interpretation of the parameter estimates is not informative, given the logit transformation of the choice dependent variable (see equations 5 and 25). We therefore provide the marginal (or partial) effects which have substantive behavioral meaning, defined as the derivatives of the choice probabilities (equation 25). A marginal effect is the influence a one unit change in an explanatory variable has on the probability of selecting a particular outcome, *ceteris paribus*[8]. The marginal effects need not have the same sign as the model parameters. Hence, the statistical significance of an estimated parameter does not imply the same significance for the marginal effect.

The generalized ordered logit model has a preferred goodness of fit over the traditional ordered logit model. With four degrees of freedom difference, the likelihood ratio of 181.92 is statistically significant on any acceptable chi-squared test level. The generalized model has included a random parameter form for congestion time framing and has accounted for two systematic sources of variation around the mean of the random threshold parameter (i.e., the accumulation of travel time and gender).

The evidence identifies a number of statistically significant influences on the number of attributes attended to, given the maximum number of attributes provided. Individuals clearly self-select attribute information to process in stated choice studies, just as they do in real markets, where the transaction costs of seeking out, compiling and assessing large amounts of potentially useful information is often seen as burdensome and/or as producing insufficient benefit. While the evidence herein cannot establish whether an attribute reduction strategy is *strictly* linked to behavioral relevance, or to a coping strategy for handling cognitive burden, both of which are legitimate paradigms in real markets, it does provide important signposts on how many attributes provided within a specific context are processed to reflect attribute relevancy.

The threshold parameter has a statistically significant mean and two sources of systematic variation across the sample around the mean threshold parameter estimate. We investigated an unconstrained random parameter normal distribution; however the standard deviation parameter estimate was not statistically significant from zero. The evidence however justifies the inclusion of a non-fixed threshold parameter, with a higher mean estimate across the sampled population when an individual aggregates the travel time components and when they are male. This is an important finding since it justifies the new formulation of the threshold parameters in ordered choice models as behaviorally meaningful. We take a closer look at each model discussing the evidence for design dimensions, framing around the base, attribute packaging, variance decomposition, and other effects. The magnitude and direction of influence is given in Table 5 for the marginal effects which have to be interpreted relative to each level of the *number of attributes ignored*.

---

[8] This holds for continuous variables only. For dummy (1,0) variables, the marginal effects are the derivatives of the probabilities given a change in the level of the dummy variable.

Table 4: Ordered logit models (2,562 observations)

Attribute	Units	Ordered Logit	Generalised Ordered Logit
Constant		2.9682 (4.17)	2.9504 (2.79)
Design Dimensions:			
Narrow attribute range	1,0	1.3738 (3.59)	1.4275 (2.35)
No. of alternatives	Number	-0.9204 (-4.1)	-1.0205 (-2.87)
Framing around Base Alt:			
Free flow time for Base minus SC alternative level	Minutes	0.0329 (4.02)	0.0599 (3.44)
Congested time for Base minus SC alternative level	Minutes	-0.0083 (-1.80)	0.0761 (2.20)
Attribute Packaging:			
Adding travel time components	1.0	-0.7407 (-4.25)	-0.8700 (-3.33)
Variance decomposition:			
Number of levels	Number	0.1043 (2.35)	0.3357 (4.48)
Free flow time for Base minus SC alternative level	Minutes	-0.0164 (-2.75)	-0.0332 (-4.04)
Who pays (1= commuter personally)	1,0	-0.3070 (5.74)	-0.3721 (-3.89)
Threshold Parameters:			
Mu1		0	0
Mu2 mean		3.0973 (5.74)	0.8753 (3.71)
Standard deviation of Mu2 threshold parameter			0.0767 (0.018)
Threshold Parameter Decomposition:			
Adding travel time components	1,0		1.7447 (10.83)
Gender (male =1)	1,0		0.3366 (2.80)
Random parameters:			
Congested time for Base minus SC alternative level	1,0		0.2652 (2.48)
Count of Choice Responses:			
	max # attributes minus #ignored		obs
0	5-0		1415
1	5-1		1080
2	5-2		66
Log-Likelihood		-1871.80	-1780.85

Table 5: Marginal effects derived from ordered logit models

Attribute	Ordered Logit	Generalised Ordered Logit
	Average No. of Attributes Ignored	Average No. of Attributes Ignored
Design Dimensions:		
Narrow attribute range	-0.4148, 0.8412, 0.0550	-0.3127, 0.2803, 0.0324
No. of alternatives	0.2779, -0.2608, -0.0171	0.2236, -0.2004, -0.032
Framing around Base Alt:		
Free flow time for Base minus SC alternative level	-0.0099, 0.0093, 0.0006	-00131, 0.0118, 0.0014
Congested time for Base minus SC alternative level	0.0025, -0.0024, -0.0002	-0.0167, 0.0149, 0.0017
Attribute Packaging:		
Adding travel time components	0.237, -.2099, -0.0137	0.1906, -0.1709, -0.0198
Variance Decomposition:		
No. of levels	-0.1104, 0.0249, 0.0856	
Free flow time for Base minus SC alternative level	-0.2386, 0.0537, 0.1849	
Who pays	0.0740, -0.0167, -0.0573	

Note: the three marginal effects per attribute refer to the levels of the dependent variable.

In commenting on the marginal effects, it should be noted that, for the generalised ordered logit model, some attributes have more than one role; for example the framing of free flow time is both a main effect influence as well as a source of variance decomposition (i.e., systematic source of heterogeneity) for the unobserved variance; and the attribute accumulation for travel time is both a main effect and a systematic source of influence on the distribution of the random threshold parameter. The generalised ordered choice model (GOCM) takes all of these sources into account in identifying the marginal effects for each level of the choice variable. In contrast, where an attribute has multiple roles in the traditional ordered choice model (TOCM), the marginal effects are calculated separately.

The dummy variable for ‘narrow attribute range’ has the greatest marginal effect, although its influence is moderated in GOCM compared to TOCM. The probability of considering all attributes from the offered set decreases as an attribute’s range narrows, *ceteris paribus*. That is, respondents tend to ignore all attributes when the difference between attribute levels is small. This result is perhaps due to the fact that evaluation of small differences is more difficult than evaluation of large differences. An important implication is that if an analyst continues to include, in model estimation, an attribute across the entire sample that is *ignored by a respondent*, then there is a much greater likelihood of mis-specified parameter estimates in circumstances where the attribute range is narrower than wider. This finding has interesting implications for the growing evidence that mean WTP for an attribute tends to be higher under a wider range for the numerator attribute (Louviere and Hensher 2001). Simply put, the greater relevance in preserving the attribute content under a wider range will mean that such an attribute is relatively more important to the outcome, than it is under a narrow range specification, and hence a higher mean WTP is inferred.

The marginal effects for the narrow attribute range are positive when one (i.e., 5-1) or two attributes (i.e., 5-2) are ignored. Importantly the positive effect is greater when one attribute is ignored than when two are ignored. This suggests that the probability of considering four or three attributes from the offered set increases as an attribute's range goes from narrow to non-narrow, *ceteris paribus*, but to a greater extent for four attributes. What we are observing across all three levels of the dependent variable is U- (or inverted U-) shaped response, which appears to be the case for all attributes in GOCM. Thus for the narrow attribute range we have the highest probability of preserving four attributes than of preserving three attributes given that the probability of preserving all attributes is decreased. Given the observed profile of the sampled respondents preserving five, four and three attributes (Table 3), where there are only 66 individuals in the last category (compared to 1415 and 1080 in 5-0 and 5-1), we have greater confidence in the relative marginal effects of preserving all (i.e., five) attributes and four attributes.

As we increase the 'number of alternatives' to evaluate (over the range of 2 to 4), *ceteris paribus*, the importance of considering all attributes increases, as a way of making it easier to differentiate between the alternatives. This finding runs counter to some views, for example, that individuals will tend to ignore increasing amounts of attribute information as the number of alternatives increases. Our evidence suggests that the processing strategy is dependent on the nature of the attribute information, and not strictly on the quantity. The negative marginal effects for ignoring one and two attributes (or preserving for and three attributes) suggest that these rules are less likely to be adopted as the number of alternatives increases.

The theoretical argument promoted in prospect theory for reference points is supported by our empirical evidence. We have framed the level of each attribute relative to that of the experienced car commute trip as (i) free flow time for current (or base) minus the level associated with an attribute and alternative in the SC design, and (ii) the congested travel time for the base minus the level associated with each SC alternative's attribute. The more that an SC attribute level deviates from the reference alternative's level, the more likely that an individual will process an increased number of attributes. This evidence was found for both the 'free flow time' and 'congested time' framing effects. Conversely, as the SC design attribute level moves closer to the reference alternative's level, individuals appear to use some approximation paradigm, in which closeness suggests similarity, and hence ease of eliminating specific attributes, because their role is limiting in differentiation.

Reference dependency not only has a direct (mean) influence on the number of attributes ignored; it also plays a role via its contribution to explaining heteroscedasticity in the variance of the unobserved effects. This has already been accounted for in the GOCM marginal effects for free flow time framing. It is separated out in the TOCM. The effect of widening the gap between the base and SC 'free flow time' reduces the heteroscedasticty of the unobserved effects across the respondents, increasing the acceptability of the constant variance condition when simpler models are specified.

In GOCM, the congested time framing effect is represented by a distribution across the sample. The random parameter has a statistically significant standard deviation parameter estimate, resulting in a distribution shown in Figure 3. The range is from -0.857 to 1.257; hence there is a sign change around the mean of 0.70833 and standard deviation of 0.2657. This results in the same mean marginal effect sign in GOCM as free flow time framing; however when we treated congested time framing as having a fixed parameters (in TOCM, where the standard deviation parameter was not statistically significant), the signs are swapped for all levels of the choice variable. The evidence from the GOCM is intuitively more plausible.

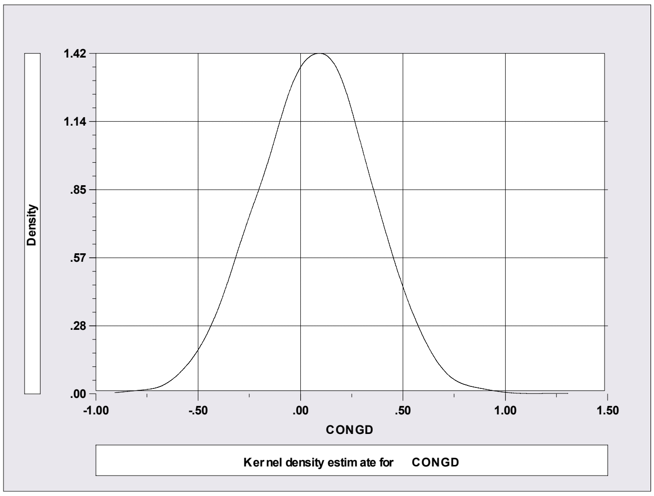

Figure 3: Distribution of preference heterogeneity for congested time framing

The attribute-accumulation rule in stage 1 editing under prospect theory is consistently strong for the aggregation of travel time components. The positive marginal effect for the dummy variable 'adding three travel time components' indicates that, on average, respondents who add up the time components, in assessing the alternatives, tend to ignore more attributes.

There is clear evidence that a relevant simplification rule is re-packaging of the attribute set, where possible, through addition. This is not a cancellation strategy, but a rational way of processing the information content of component attributes, and then weighting this information (in some unobserved way) in comparing alternatives.

The socio-economic characteristics of respondent's proxy for other excluded contextual influences. A respondent's role in paying the toll was identified, through its influence on variance decomposition of the unobserved effects, as a statistically significant socio-economic influence on the number of attributes considered. We have no priors on the likely sign of the influence on variance. The positive marginal effect for who pays suggests that those who pay themselves tend to increase the variance of the unobserved effects, resulting in a lower probability of preserving more attributes. Gender was a systematic source of influence on the threshold parameter, increasing its mean estimate for males.

5. Conclusions

The recognition of randomness in the threshold parameters and the identification of systematic sources of heterogeneity in the mean threshold parameter estimate is an important extension of the existing ordered choice model. This paper has brought together all of the contributions in the literature and extended them, in particular to ensure preservation of the ordering of thresholds in the context of random parameterisation of the thresholds (equations 16 to 21). The specific application herein, on the role that attributes play in choice making in stated choice experiments, pivoted around a real market experience, has highlighted the role of random thresholds and decomposition, suggesting that the generalized empirical model is a rich behavioral addition to the literature on ordered choice modeling.

References

Allison, P. (1999): 'Comparing Logit and Probit Coefficients across Groups', *Sociological Methods and Research* 28, 186-208.

Arentze, T., Borgers, A., Timmermans, H. and DelMistro, R. (2003): 'Transport Stated Choice Responses: Effects of Task Complexity, Presentation Format and Literacy', *Transportation Research*, 39E, 229-244.

Rose, J.M., Bliemer, M.C., Hensher and Collins, A. T. (2008): 'Designing Efficient Stated Choice Experiments in the Presence of Reference Alternatives', *Transportation Research B* 42 (4), 395-406

Boes, S., and R. Winkelmann (2004): Income and Happiness: New Results from Generalized Threshold and Sequential Models. *IZA Discussion Paper No. 1175*, SOI Working Paper No. 0407.

Boes, S. and Winkelman, R. (2007): ‚Ordered Response Models', Boes, S. and R. Winkelmann, *Allgemeines Statistiches Archiv Physica Verlag*, 90(1), 165-180.

Bonini, N., Tentori, K. and Rumiati, R. (2004): 'Contingent Application of The Cancellation Editing Operation: The Role of Semantic Relatedness Between Risky Outcomes', *Journal of Behavioral Decision Making*, 17, 139-152.

Brant, R. (1990): 'Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression', *Biometrics*, 46, pp. 1171-1178.

Brewer, C., Kovner, Wu, Y, Greene, W., Liu, Y. and Reimers, C. (2006): 'Factors Influencing Female Registered Nurses' Work Behavior,' *Health Services Research*, 41, 3, 860-866.

Campbell, D., Hutchinson, W. and Scarpa, R. (2008, in press): 'Incorporating Discontinuous Preferences into the Analysis of Discrete Choice Experiments' *Environment and Resource Economics*,

Cantillo, V., Heydecker, B. and Ortuzar, J. de Dios (2006): 'A Discrete Choice Model Incorporating Thresholds for Perception in Attribute Values', *Transportation Research* B, 40 (9), 807-825.

Chamberlain, G. (1980): 'Analysis of Covariance with Qualitative Data', *Review of Economic Studies*, 47, 225-238.

Daykin,A. and Moffitt, P. (2002): 'Analyzing Ordered Responses: A Review of the Ordered Probit Model', *Understanding Statistics*, 3, 157-166.

De Shazo, J.R. and Fermo, G. (2002): Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *Journal of Environmental Economics and Management* 44, 123-143.

Eluru, N., Bhat, C. R. and Hensher, D.A. (2008): 'A Mixed Generalized Ordered Response Model for Examining Pedestrian and Bicyclist Injury Severity Level in Traffic Crashes', Accepted for *Accident Analysis and Prevention*

Greene, W.H. (1997): LIMDEP Version 7.0 Reference manual, Econometric Software.

Greene, W.H. (2002): LIMDEP Version 8.0 Reference manual, Econometric Software.

Greene, W.H. (2007): NLOGIT Version 4.0 Reference manual, Econometric Software.

Greene, W.H. (2008): *Econometric Analysis, 6th Ed.*, Prentice Hall.

Greene, W. H. Harris, M., Hollingsworth, B. and Maitra, P. (2008): A Bivariate Latent Class Correlated Generalized Ordered Probit Model with an Application to Modeling Observed Obesity Levels, *Working Paper 08-18*, Stern School of Business, New York University.

Hallahan, K. (1999): 'Seven Models of Framing: Implications for Public Relations', *Journal of Public Relations Research*, 11(3), 205-242.

Hensher, D.A. 2001: 'Measurement of the Valuation of Travel Time Savings', *Journal of Transport Economics and Policy* 35(1):71–98.

Hensher, D.A. (2004): 'Accounting for Stated Choice Design Dimensionality in Willingness to Pay for Travel Time Savings', *Journal of Transport Economics and Policy*, 38 (2), 425-446.

Hensher, D.A. (2006a): 'Revealing Differences in Behavioral Response due to the Dimensionality of Stated Choice Designs: an Initial Assessment', *Environmental and Resource Economics* 34 (1), May, 7-44.

Hensher, D.A. (2006b): 'How Do Respondents Handle Stated Choice Experiments? - Attribute Processing Strategies under Varying Information Load', *Journal of Applied Econometrics,* 21, 861-878

Hensher, D.A. (2008): 'Joint Estimation of Process and Outcome in Choice Experiments and Implications for Willingness to Pay', *Journal of Transport Economics and Policy*, 42 (2), May, 297-322.

Hensher, D.A. and Greene, W.H. (2003): 'Mixed Logit Models: State of Practice', *Transportation*, 30 (2), 133-176.

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

Hensher, D.A., Rose, J. and Greene, W. (2005a): 'The implications on willingness to pay of respondents ignoring specific attributes', *Transportation*, 32(2), 203-222.

Hess, S., Rose, J. M. and Hensher, D.A. (2008): 'Asymmetrical Preference Formation in Willingness to Pay Estimates in Discrete Choice Models', *Transportation Research E.*

Houston, D. A., Sherman, S. J. and Baker, S. M. (1989): 'The Influence of Unique Features and Direction of Comparison on Preferences', *Journal of Experimental Social Psychology*, 25, 121–141.

Houston, D. A. and Sherman, S. J. (1995): 'Cancellation and Focus: The Role of Shared and Unique Features in the Choice Process', *Journal of Experimental Social Psychology*, 31, 357–378.

Kahnemann, D. and Tversky, A. (1979): 'Prospect Theory: An Analysis of Decisions under Risk', *Econometrica*, 47 (2), 263-91.

Long (1997), *Regression Models for Categorical and Limited Dependent Variables*, Sage, New York.

Long and Frees (2006) *Regression Models for Categorical and Limited Dependent Variables Using Stata*, Stata Press.

Louviere, J.J. and Hensher, D.A. (2001): 'Combining Preference Data', in Hensher, D.A. (ed.) *The Leading Edge of Travel Behaviour Research*, Pergamon Press, Oxford, 125-144.

Louviere, J.J., Hensher, D.A. and Swait, J.F. (2000) *Stated Choice Methods and Analysis*, Cambridge University Press, Cambridge.

Mundlak (1978): 'On the Pooling of Time Series and Cross Sectional Data', *Econometrica,* 56, pp. 69-86.

Pratt, J. (1981): 'Concavity of the Log Likelihood', *Journal of the American Statistical Association*, 76, 103-106.

Pudney, S. and Shields, M. (2000): 'Gender, Race, Pay and Promotion in the British Nursing Profession: Estimation of a Generalized Ordered Probit Model', *Journal of Applied Econometrics*, 15 (4), 367-399.

Rose, J.M. and M.C.J. Bliemer (2007): 'Stated Preference Experimental Design Strategies' in Hensher, D.A. and K. Button, (eds.) *Transport Modelling*, Second Edition, Handbooks in Transport, Vol. 1, Elsevier Science, Oxford, Chapter 8.

Starmer, C. (2000): 'Developments in Non-expected Utility Theory: the Hunt for a Descriptive Theory of Choice under Risk', *Journal of Economic Literature*, XXXVIII, 332-382.

Starmer, C. and Sugden, R. (1993): 'Testing for Juxtaposition and Event Splitting Effects', *Journal of Risk and Uncertainty*, 6, 235-54.

Stewart, M.B. (2004): A Comparison of Semiparametric Estimators for the Ordered Response Model', *Computational Statistics and Data Analysis* 49, 555–573.

Swait, J. (2001): 'A Non-Compensatory Choice Model Incorporating Attribute Cut-Offs', *Transportation Research* B, 35(10), 903-928.

Swait, J. and Adamowicz, W. (2001a): 'The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching', *Journal of Consumer Research,* 28, 135-148.

Swait, J. and Adamowicz, W. (2001b): 'Choice Environment, Market Complexity, and Consumer Behavior: A Theoretical and Empirical Approach for Incorporating Decision Complexity into Models of Consumer Choice', *Organizational Behavior and Human Decision Processes,* 49, 1-27.

Terza, J. (1985) 'Ordinal Probit: A Generalisation', *Commmunications in Statistics – Theory and Methods*, 14 (1), 1-11.

Train, K. and Wilson, W.W. (2008): 'Estimation on Stated Preference Experiments Constructed from Revealed Preference Choices', *Transportation Research B*, 42 (3), 191-203.

Thaler, R. (1999): 'Mental Accounting Matters', *Journal of Behavioral Decision Making*,

Williams, R. (2006): 'Generalized Ordered Logit/ Partial Proportional Odds Models for Ordinal Dependent Variables', *The Stata Journal* 6 (1), 58-82.

Zavoina, R. and McElvey, W. (1975): 'A Statistical Model for the Analysis of Ordinal Level Dependent Variables', *Journal of Mathematical Sociology*, Summer, 103-120.

Appendix A

Designs for five-attributes

Block	Scenarios	Alternative 1					Alternative 2				
		Free Flow time	Slowed down time	Stop/Start time	Uncert of travel time	Total cost	Free Flow time	Slowed down time	Stop/Start time	Uncert of travel time	Total cost
1	1	1	1	0	1	1	0	1	0	0	1
1	2	0	1	0	1	0	1	0	1	1	0
1	3	1	1	0	1	1	0	1	0	1	1
1	4	0	0	1	1	0	1	1	0	0	1
1	5	0	1	0	0	1	1	0	1	1	1
1	6	0	0	0	1	1	0	0	0	0	1
1	7	0	1	1	1	1	1	0	0	0	1
1	8	1	1	1	1	0	0	1	0	1	1
1	9	0	0	1	1	0	0	1	1	0	1
1	10	0	0	1	0	0	0	0	0	0	0
1	11	0	1	1	1	1	0	0	1	0	1
1	12	0	0	1	0	0	1	0	0	0	0
1	13	1	1	1	0	1	0	0	1	1	1
1	14	1	1	0	0	1	0	0	1	1	0
1	15	1	0	0	0	1	1	1	1	0	1
2	1	0	0	1	0	0	0	0	0	1	0
2	2	1	1	1	1	0	0	1	0	0	0
2	3	1	1	1	1	1	1	1	1	1	0
2	4	0	1	1	0	0	1	0	1	0	0
2	5	1	0	0	0	1	1	0	1	1	0
2	6	1	0	0	0	1	1	0	0	0	0
2	7	1	1	0	0	0	0	0	1	1	0
2	8	0	1	0	0	1	1	1	1	0	1
2	9	1	1	1	1	0	0	0	0	0	1

2	10	1	0	1	0	0	0	1	0	1	0
2	11	1	1	0	1	1	0	0	1	1	1
2	12	1	0	0	1	0	1	0	0	1	1
2	13	1	0	1	0	0	0	1	0	1	0
2	14	0	1	1	1	0	1	1	1	1	1
2	15	1	0	0	0	0	0	1	0	0	0

Block	Scenarios	Alternative 1					Alternative 2				
		Free Flow time	Slowed down time	Stop/Start time	Uncertainty of travel time	Total cost	Free Flow time	Slowed down time	Stop/Start time	Uncertainty of travel time	Total cost
1	1	1	0	0	3	3	3	3	1	1	1
1	2	2	0	3	2	1	0	1	0	1	2
1	3	1	1	3	2	2	3	0	1	0	3
1	4	2	3	1	3	2	0	1	2	2	3
1	5	2	1	2	1	1	1	3	0	0	0
1	6	2	1	1	0	0	3	2	0	3	1
1	7	3	0	2	1	3	2	2	3	3	2
1	8	0	3	2	0	1	3	2	1	2	0
1	9	0	0	3	1	1	3	3	2	2	0
2	1	1	3	3	1	3	2	2	2	0	2
2	2	0	0	0	2	2	1	2	2	1	0
2	3	0	3	3	3	0	1	1	0	0	1
2	4	1	1	1	3	1	2	2	0	1	0
2	5	2	3	0	2	3	1	2	3	0	3
2	6	0	2	1	3	3	3	3	3	0	2
2	7	3	1	3	0	0	1	0	1	1	2
2	8	2	1	0	3	3	0	0	2	2	2
2	9	0	2	1	2	1	3	0	2	3	0

Block	Scenarios	Alternative 1					Alternative 2					Alternative 3				
		Free Flow time	Slowed down time	Stop/Start time	Uncertainty of travel time	Total cost	Free Flow time	Slowed down time	Stop/Start time	Uncertainty of travel time	Total cost	Free Flow time	Slowed down time	Stop/Start time	Uncertainty of travel time	Total cost
1	1	2	3	3	3	1	3	1	1	0	2	1	0	0	2	3
1	2	0	2	0	2	1	2	0	2	1	0	3	1	3	3	0
1	3	2	0	1	3	2	0	1	2	0	1	1	3	3	1	0
1	4	0	3	1	3	0	3	2	0	1	0	2	1	3	0	1
1	5	3	2	2	3	2	0	3	3	0	3	1	1	1	2	3
1	6	2	2	1	0	0	0	0	0	1	2	1	3	2	2	1
2	1	0	1	0	3	3	2	2	1	2	3	1	3	2	0	2
2	2	3	1	2	2	0	0	2	3	1	2	1	0	0	3	1
2	3	3	3	0	0	0	1	1	1	1	1	0	0	3	2	2
2	4	2	3	0	1	3	3	0	3	2	1	1	2	2	3	2
2	5	2	0	2	3	3	3	3	1	1	1	1	2	3	0	0
2	6	2	1	0	2	2	3	0	1	0	3	0	2	2	1	3

Block	Scenarios	Alternative 1					Alternative 2					Alternative 3					Alternative 4				
		Free Flow time	Slowed down time	Stop/Start time	Uncertainty of travel time	Total cost	Free Flow time	Slowed down time	Stop/Start time	Uncert of travel time	Total cost	Free Flow time	Slowed down time	Stop/Start time	Uncert of travel time	Total cost	Free Flow time	Slowed down time	Stop/Start time	Uncert of travel time	Total cost
1	1	2	1	0	0	2	0	2	0	0	2	1	0	1	1	1	0	1	2	2	2
1	2	1	0	0	0	1	2	1	2	2	0	0	2	1	1	2	0	2	1	1	2
1	3	1	2	2	0	0	2	0	0	1	0	0	2	2	2	2	0	1	1	2	1
1	4	1	2	2	0	1	0	0	1	2	2	1	2	2	2	2	2	1	0	1	1
1	5	2	2	2	1	1	1	1	1	0	1	2	2	2	1	2	0	0	0	2	2
1	6	2	2	0	2	1	1	0	1	0	2	1	0	1	0	2	0	1	2	1	1
1	7	0	0	0	1	2	1	2	2	2	1	2	1	1	0	2	2	1	1	0	2
1	8	0	1	1	1	0	1	2	0	2	0	2	0	2	0	1	2	0	2	0	1
1	9	1	1	0	2	2	2	2	2	1	0	1	1	0	2	2	0	0	1	0	1
1	10	2	2	1	2	1	0	0	2	0	0	1	1	0	1	0	2	2	1	2	1
1	11	1	2	1	0	0	2	0	0	1	1	0	1	2	2	0	2	0	0	0	2
1	12	0	2	0	0	0	2	1	1	2	0	1	0	2	1	1	1	0	2	1	1
2	1	2	0	1	1	2	2	0	1	1	1	0	2	2	2	0	1	1	0	0	0
2	2	0	0	1	2	1	2	2	2	0	0	1	1	0	1	2	1	1	0	1	2
2	3	0	1	1	0	2	2	0	2	1	0	1	2	0	2	0	0	1	1	0	1
2	4	2	2	1	2	0	1	0	2	1	0	0	1	0	0	1	0	1	0	0	2
2	5	1	2	2	2	2	1	2	1	1	0	0	1	2	0	0	2	0	0	2	1
2	6	1	1	0	2	1	0	0	2	1	2	0	0	2	1	2	2	2	1	0	0
2	7	1	1	2	2	0	0	2	1	1	1	2	0	0	0	0	0	2	1	1	1
2	8	1	0	1	2	2	2	1	2	1	0	1	0	1	2	2	0	2	0	0	1
2	9	1	2	0	1	1	0	1	2	0	1	2	0	1	2	0	0	1	0	1	2
2	10	2	1	1	0	0	1	0	2	2	2	1	0	2	2	1	0	2	0	1	0
2	11	2	2	0	0	2	0	0	0	2	1	2	2	2	0	0	1	1	1	1	0
2	12	1	0	2	0	0	0	2	1	1	0	2	1	0	2	2	2	1	0	2	1