Embedding Decision Heuristics in Discrete Choice Models: A Review

Leung, W. And Hensher ,D.A.

ABSTRACT

Contrary to the usual assumption of fixed, well-defined preferences, it is increasingly evident that individuals are likely to approach a choice task using rules and decision heuristics that are dependent on the choice environment. More specifically, heuristics that are defined by the local choice context, such as the gains or losses of an attribute value relative to the other attributes, seem to be consistently employed. Recent empirical findings also demonstrate that previous choices and previously encountered choice tasks shown to respondents can affect the current choice outcome, indicating a form of inter-dependence across choice sets.

This paper is primarily focused on reviewing how heuristics have been modelled in stated choice data. The paper begins with a review of the heuristics that may be relevant for coping with choice task complexity and then proceeds to discuss some modelling approaches. Next, relational heuristics, such as prospect theory, random regret minimisation and extremeness aversion (compromise effect) are discussed. These are heuristics which operate within the local choice set. Another major class of heuristics reviewed in this paper pertains to ordering effects and more generally, on past outcomes and past attribute levels of the alternatives.

Keywords: heuristics; thresholds; contextual effects; relational heuristics; compromise effect

Introduction

The typical approach used in much of the discrete choice modelling literature in traveller behaviour studies assumes that well-defined preferences exist for most decision tasks. Under the standard random utility theory, preferences are stable and invariant to choice tasks, and are fully known to the respondent. In the great majority of cases, the analyst writes out a utility function assuming that the respondent is cognitively indefatigable, examining all alternatives and all attributes across all choice tasks in the same fully compensatory manner. The linear weighted additive form for utility, estimated by means of the multinomial logit model, has been found to be a convenient representation which is capable of embodying all these assumptions, and has therefore become the mainstay in discrete choice modelling.

As a description of how people behave, research in the decision behaviour field, and more recently from the choice experiment literature, has cast considerable doubt that the weighted additive function, assumed independent of contextual effects, comes close to being an accurate representation of the actual processes used in the majority of decision tasks (see for example, McFadden, 1999; Gigerenzer *et al.*, 1999). Empirical research from the psychology literature has shown that preferences for an alternative are influenced by the choice context itself, in other words, by factors that are beyond the immediate attributes of the alternative under consideration. Earlier work focused on how choice task properties, such as the number of alternatives and attributes, impact decisions in terms of how decision rules are selected and applied (e.g., Payne *et al.*, 1993). Decisions may also be made according to some reference point selected by the respondent (see for example, Kahneman and Tversky, 1979; Tversky and Kahneman, 1991; Li and Hensher, 2011). This reference may have something to do with the other alternatives in the same choice set or even across previously encountered choice sets (Simonson and Tversky, 1992; Kivetz *et al.*, 2004). Moreover, different respondents may be attending to various subsets of attributes, and such heterogeneity may be masked if it is assumed that preference weights are the same across the entire dataset. Outside the immediate environment of a choice experiment, the social context, through an intermediate construct of comparative happiness, may also influence utility (Abou-Zeid and Ben-Akiva, 2011). Compared to the context independent, linear weighted additive form of utility, these studies suggest that other representations of utility which better approximate the realism of real-life decision making can lead to better goodness of fit and more plausible model estimates and outputs.

The purpose of this paper is to review the decision heuristics that have been modelled in the discrete choice literature, drawing from extant work in the fields of transport, environment, marketing and psychology. The examples cited in this paper primarily deal with stated choice data. Heuristics prompted by contextual effects in the form of choice task properties are first discussed, followed by a review of the various methodological approaches to embed such heuristics into choice models. Contextual effects embodied in relational heuristics linked to the idea of referencing are then discussed. The paper concludes with some suggestions for the future direction of this research.

Contextual Effects Arising from Choice Task Properties

Heuristics and choice task properties

The psychology literature has amassed a wealth of evidence to suggest that humans rely on the use of quick mental processing rules known as heuristics to manage the vast number of decisions that must be made in everyday life. While the fully compensatory weighted additive rule is commonly used for the purposes of modelling, psychologists have argued that this rule implies an assumption of stable, well-articulated preferences which appears to hold only under conditions where the choice task is familiar or when the respondent has experience with the various alternatives that are presented (Payne *et al.*, 1993). In many instances, it is argued that these conditions fail to apply, and preferences are not determined in advance of the choice situation, but are instead constructed in response to the contextual effects which vary according to the properties of the choice task. As described by Payne *et al.* (1999, p. 245), the construction process involves an interaction between "the properties of the human information processing system and the properties of the choice task."

Therefore, rather than static decision processes which are repeatedly applied to different choice contexts, the conclusion drawn by behavioural decision research is that "individuals have a repertoire of decision strategies for solving decision problems" (Bettman *et al.*, 1998, p. 194). Some decision strategies that might be a part of this repertoire include satisficing (Simon, 1955), lexicography (Tversky, 1969), elimination-by-aspects (EBA) (Tversky, 1972) and the majority of confirming dimensions (Russo and Dosher, 1983). Brief descriptions of these heuristics follow:

(1)	*Satisficing*: Under this heuristic, an attribute may be associated with a pre-defined cut-off and the first alternative with all attribute levels satisfying the cut-off criteria is chosen. If none of the alternatives meets the cut-off criteria, the cut-offs may be relaxed and the process repeated; or a random choice is made.

(2)	*Lexicography*: The respondent evaluates all alternatives based on what is deemed to be the most important attribute. The alternative with the best value on that attribute is chosen. If there is a tie, or if the difference between the best levels of the attribute is not noticeable, then the remaining alternatives are evaluated on the next most important attribute, and so on.

(3)	*EBA*: The respondent identifies the most important attribute (either deterministically or probabilistically) and its associated cut-off threshold. An alternative is eliminated if its attribute fails to satisfy the cut-off. This process is repeated with the second most important attribute and so on until one alternative remains.

(4)	*Majority of confirming dimensions*: The first two alternatives are compared and the one with the larger number of winning attributes is retained. The retained alternative is compared with the next alternative and so on until all alternatives have been evaluated. The alternative with the highest number of winning attributes is selected.

Payne *et al.* (1993) argue that heuristics are used to manage situations of high choice task complexity. Choice complexity is largely determined by the choice context, which in this case are

3

choice task properties such as the number of alternatives in the choice set, the number of attributes in each alternative and the correlation between attribute levels across multiple alternatives. This idea of complexity can be contrasted with Hensher's (2006) notion of relevancy, which pertains to providing more complete descriptions of attributes in the choice task and allowing respondents to form their own processing rules with regards to relevancy. Hence, a choice task that disaggregates say a time attribute into its various components such as free-flow and slowed down time may be more relevant to a respondent, even though this task would be "more complex" from Payne *et al.*'s (1993) perspective.

An effort-accuracy trade-off framework has been proposed as one possible mechanism by which individuals select the decision strategy or combination of decision strategies out of the repertoire available to them (Payne *et al.*, 1993). This framework is not a formal choice model, but it still postulates a relationship between decision strategies and choice task complexity. Essentially, the choice of heuristic is thought to be the outcome of trading-off between two conflicting goals: maximising the accuracy of a decision, with the weighted linear additive rule defining the normative benchmark level of accuracy, and minimising the cognitive effort required to reach that decision.

Using qualitative evidence, Payne *et al.* (1993) suggest that heuristics such as lexicography and EBA are more commonly used when choice tasks become more complex. These heuristics are not much less accurate compared to the weighted additive rule, but they require less cognitive resources. The choice of heuristic is not necessarily only as a strict consequence of the dimensionality of a choice experiment, but can be associated with previous experience in adopting specific heuristics, and hence it is important to recognise and account for both potential sources of influence on choice making and selection of a heuristic.

Conceptually, the effort-accuracy framework requires the decision maker to be cognisant of the costs and benefits of each strategy as applied to the choice task under consideration. The realism of such an assumption is debatable, in view of the maintained hypothesis that cognitive effort is a scarce resource. At the empirical level, a significant challenge remains in terms of identifying and quantifying the cognitive effort associated with each heuristic. Nonetheless, the effort-accuracy framework describes broad conditions under which non-compensatory heuristics like lexicography or EBA are more likely to be used, thus providing some guidance in terms of model specification. For example, Young (1984) assumes an EBA model of choice because of the complexity of choosing a residential location among many attributes and alternatives.

Discrete choice models of information processing strategies such as attribute non-attendance (see for example, Hensher, 2010; Hensher and Greene, 2010; Scarpa *et al.*, 2009) may be viewed as a "bottom-up" or "data-driven" view of preference construction (Payne *et al.*, 1993, p. 171). Respondents can shape or change decision strategies by exploiting previously encountered problem structures. Decision problems are subsequently restructured as an intermediate step, making them more amenable to analysis using certain heuristics. Information in choice tasks might be transformed through rounding or standardising values in a common metric. Information might also be rearranged or further simplified by deeming certain attributes irrelevant. It is argued that

4

such restructuring serves to reduce the perceived complexity of the choice task (Payne and Bettman, 1992).

Modelling perspectives on choice task complexity

As a corollary of the effort-accuracy trade-off framework, a phased decision strategy may also be employed under some circumstances (Stevenson *et al.*, 1990). In this strategy, respondents are thought to initially rely on some non fully-compensatory heuristic to reduce choice task complexity before using a fully compensatory strategy to evaluate the reduced number of alternatives and/or attributes. This type of decision making can be incorporated into a discrete choice model through a two stage decision process, with screening rules invoked in the first stage to select a subset of alternatives from a larger universal set. The final choice is made from the reduced set. The screening rules could be based on the history of past choices, or on the attribute levels of alternatives in the current choice situation. In these models, Swait and Ben-Akiva (1987) have observed that the first stage resembles the "considered subset" in Simon's (1955) satisficing model. A general form of a two-stage model, attributed to Manski (1977), is given in Equation (1):

$$P_j = \sum_{C \in G} P(j \mid C) P(C)$$

(1)

where P_j is the unconditional probability that alternative *j* is chosen, $P(j \mid C)$ is the probability of choosing alternative *j* given the reduced choice set C, and $P(C)$ is the probability that the reduced choice set is *C*, among all the non-empty subsets of a master choice set *G*.

Expanding on the Manski equation, Cantillo and Ortuzar (2005) assume a first stage elimination involving the use of a rejection mechanism based on individual-specific thresholds of attribute levels. Alternatives which survive the first stage screening are then evaluated in the usual compensatory manner within the random utility framework. Cantillo and Ortuzar suggest that the threshold might be determined by "the most favourable value among those that the attribute can take for the set of potential alternatives; it could also be the value that the attribute takes for the chosen alternative or simply any reference value" (Cantillo and Ortuzar, 2005, p. 644). Hence, there is a great amount of flexibility as to how the individual-specific thresholds are modelled. In general, the vector of thresholds may be distributed across all individuals with a certain mean and variance-covariance structure with this distribution possibly being made a function of individual specific characteristics. In a stated route-choice experiment of possible car trips, Cantillo and Ortuzar find evidence for an age varying threshold effect for the accident rate attribute but no such effect for the attributes of time and cost. However, in their model, the situation where a very low attribute level for cost or time might invoke a believability threshold is not considered.

Likewise, in Suzuki (2007)'s two-step model of airline and airport choice, the probability of an alternative surviving the first stage is assumed to be negatively related to its likelihood of

5

violating some minimum acceptable standard or threshold. This threshold is unknown for each respondent and is arbitrarily approximated by appealing to contextual effects using a decision rule reminiscent of lexicography and majority of confirming dimensions. The decision rule embeds various assumed threshold specifications, with the probability of survival depending on the number of attributes of an alternative being, say, the 'best' or 'best'/ 'second-best' performer in the choice set. Suzuki (2007) finds that the goodness of fit measures support the use of such thresholds in a two-step model of airline choice over conventional specifications of the value function which assume no such thresholds.

Another application of two-stage models lies in modelling endogenous attribute attendance and attribute non-attendance (Hole, 2011; Greene and Hensher, 2012). In the first stage, respondents are assumed to decide which subset of attributes to take into account. Subsequently, in the second stage, all available alternatives are evaluated conditioned on the subset of attributes chosen in the first stage. The probability of attending to a certain attribute may also be specified as a function of individual level observed characteristics. Hole's results conclusively reject the standard fully compensatory logit model in favour of the two stage model. However, this approach seems to be computationally feasible only when the number of attributes is small.

Instead of assuming a 'hard' cut-off constraint, Swait (2001) allows cut-offs to be violated, but at a cost to overall utility through a larger penalty in the utility function. An alternative whose attributes violate the cut-off can still be chosen provided sufficient compensation in the other attributes is available. This model relies on obtaining cut-off information from respondents directly. Estimation results show that the inclusion of penalty parameters is able to significantly improve the goodness-of-fit of the model. However, for empirical identification of the penalty parameters, sufficient choices must be made under conditions where cut-offs are violated and this may be easier to achieve with experimentally manipulated stated choice data rather than revealed preference data. Hensher and Rose (2011) extend Swait's model by allowing the entire utility function, including the penalty function, to be conditioned on individual specific perceptions, such as the acceptability of the alternative, the certainty of the choice response and the incidence of attribute levels in the perceived attribute threshold rejection region. Besides improved fit and in sample prediction success, conditioning on these perceptions, especially on alternative acceptability, leads to noticeable differences in mean direct elasticities compared to a model without such conditioning.

Another approach, used by Swait (2009), to formalise decision process heterogeneity is to consider a mixed model of random utility, where an alternative may be evaluated in one of several discrete states, with each state corresponding to a different decision rule or cognitive process. One of these states pertains to the usual utility maximising, fully compensatory condition, while other states may represent a more extreme version of attractiveness or unattractiveness, which aims to capture the possible use of a non-compensatory strategy, context dependence and/or attribute independence. Equation (2) illustrates this model for a simple two-condition scenario, where alternative j is assigned to either the first state which represents the trade-off condition in the usual sense, or another state representing a rejection condition, where the utility for alternative j is not defined over attribute values.

$$U_j \begin{cases} = V_j + \varepsilon_j & \text{with probability } p_j \\ = -\infty & \text{with probability } q_j \end{cases}$$
(2)

Swait's (2009) model can be set up to embed the EBA heuristic as part of choice set formation, by allowing q_j, the probability of an alternative being in the rejection condition, to be written as a function of a disjunctive screening rule: it takes just one attribute to fail the threshold cut-off before the alternative is eliminated. Conversely, p_j, which is the probability that an alternative is in the usual random utility maximising, fully compensatory trade-off condition, is written in the conjunctive sense: it is the probability of all attributes satisfying the threshold criteria before fully compensatory processing takes place.

In the model, for each attribute of interest, individual-specific thresholds can be assumed to be randomly distributed across the population, according to say, a normal distribution with mean $\overline{\tau}_k$ and variance σ_k^2. Consider an example where the EBA heuristic is applied to one aspect, for example, departure time. Then Equation (3) is obtained, where τ_k takes on a lower bound threshold (i.e., departure time no earlier than τ_k):

$$p_j = \Pr(\tau_k < X_{jk}) = \Pr\left(Z < \frac{X_{jk} - \overline{\tau}_k}{\sigma_k} \right) = \Phi\left(\frac{X_{jk} - \overline{\tau}_k}{\sigma_k} \right)$$

$$q_j = 1 - p_j = 1 - \Phi\left(\frac{X_{jk} - \overline{\tau}_k}{\sigma_k} \right)$$
(3)

Other parameterisations of p_j are also possible, for example, as a logistic function of attributes or person characteristics. Using such a logistic specification, Swait (2009) concludes that these mixed models are preferred over the standard linear additive model despite the increase in the number of parameters estimated. Andersen *et al.* (2007) use a conceptually similar mixture model to represent dual latent decision processes in decision making, with the result that respondents appear to assign a nearly two-third probability weight to an aspiration-type process that incorporates thresholds, and assigning only a one-third weight to an expected utility process.

In a more direct approach of embedding complexity into discrete choice models, Swait and Adamowicz (2001) suggest that a formal relationship exists between the error variance or scale in preferences, and entropy, which can be seen as a measure of complexity. Entropy is defined in Equation (4) with π_{js} denoting the probability of choosing alternative *j* in choice situation *s* and π_{js} obtained by *a priori* estimating a basic MNL model:

$$H_s = -\sum_{j \in S} \pi_{js} \log(\pi_{js})$$
(4)

7

As preferences among alternatives become more indistinguishable, π_{js} approaches 1/J and a high level of entropy (and complexity) is obtained. Entropy is related to scale by noting that at low and high levels of entropy, the scale is high as decision-making is relatively easy in the former case and alternatives are all approximately similar in utility terms in the latter. At moderate levels of complexity however, more preference inconsistency (lower scale) may be evident as respondents resort to using simplifying heuristics. Hence, the scale of choice task s, μ_s, may be related to H_s through a quadratic form to account for these non-linear effects, as in Equation (5):

$$\mu_s = \exp(\theta_1 H_s + \theta_2 H_s^2) \tag{5}$$

These theoretical predictions are confirmed by the empirical results in the majority of cases analysed by Swait and Adamowicz (2001). More recently, Zhang and Adamowicz (2011) show that directly embedding entropy and other contextual attributes into the preference and scale functions of a discrete choice model can control for the effects of different choice formats on preference elicitation.

Relational Heuristics

Prospect Theory

Another broad category of heuristics might be called "relational" heuristics. Unlike the models discussed earlier, choice task properties and complexity are not the focus of attention here. Instead, these heuristics emphasise the use of some reference point(s), or in some specific applications, a comparison of ratings of one alternative against another. Prospect theory (Kahneman and Tversky, 1979) is perhaps one of the most well-known behavioural model in this category. Some brief remarks about prospect theory follow, but for a more comprehensive treatment, see Van de Kaa (2010a; 2010b) and Li and Hensher (2011).

Essentially, prospect theory assumes that respondents frame alternatives and attributes relative to a reference state, that respondents are loss averse, and that there is diminishing sensitivity to gains and losses. Van de Kaa (2010a)'s meta-analysis shows that employing the assumptions of prospect theory rather than utility theory results in improved descriptive ability of travel behaviour in the vast majority of studies reviewed. Avineri and Prashker (2004) conclude that traveller behaviour in the light of travel time uncertainties violate the assumptions of expected utility theory and respondents are possibly prospect maximisers instead, while Senbil and Kitamura (2004) observe reference dependency in the decision frames relating to departure time decisions. On a more cautionary note however, Timmermans (2010) raises some questions regarding the application of prospect theory to travel behaviour under uncertainty, especially in dynamic situations where there is feedback.

The Compromise Effect and Regret Minimisation

More specific examples of relational heuristics abound. We focus the rest of the discussion on models of extremeness aversion (Simonson and Tversky, 1992; Tversky and Simonson, 1993) and regret minimisation (Chorus *et al.*, 2008; Chorus, 2010).

The extremeness aversion heuristic has been put forward as a possible explanation for the fairly robust empirical findings of the so-called "compromise effect", which leads respondents to prefer an in-between alternative when extreme alternatives are available in the choice set. Extreme alternatives are those which perform best on some attributes, but worst on others. Loss aversion may explain the compromise effect; in that the disadvantages of an alternative (defined relative to the other alternatives in the choice set) are weighted more heavily than its advantages (Tversky and Kahneman, 1991; Simonson and Tversky, 1992). Hence, the in-between alternative, with its smaller advantages and disadvantages, is more highly favoured compared to the extreme alternative. Louviere and Myer (2008) also show that this effect can arise from preference uncertainty among risk averse individuals.

The random regret minimisation (RRM) model, proposed in Chorus *et al.* (2008) and subsequently refined by Chorus (2010) has been shown to be able to accommodate the compromise effect. Regret (Bell, 1982; Loomes and Sugden, 1982) is said to occur when a non-chosen alternative leads to a more desirable outcome, for example, when a foregone alternative performs better on a certain attribute compared to the chosen alternative. Under the RRM, respondents are assumed to engage in regret avoidance behaviour by choosing the alternative which minimises regret. The regret for any considered alternative j, denoted $Reg(j)$, is the sum of all binary regrets of choosing alternative j over the non-considered alternatives $j' \in s$. Specifically:

$$Reg(j) = \sum_{\substack{j' \neq j, \\ j' \in s}} Reg(j, j') = \sum_{\substack{j' \neq j, \\ j' \in s}} \sum_{k} \ln\left(1 + \exp\left[\beta_k (X_{j'k} - X_{jk})\right]\right) \tag{6}$$

In the limit as $\beta_k (X_{j'k} - X_{jk})$ becomes sufficiently negative, $Reg(j, j')$ with respect to attribute k falls towards zero. Likewise, if $\beta_k (X_{j'k} - X_{jk})$ becomes sufficiently large, $Reg(j, j')$ with respect to attribute k approaches $\beta_k (X_{j'k} - X_{jk})$.

The RRM model is said to be semi-compensatory in the sense that improvements or deteriorations in an alternative depends very much on the relative attribute level compared to other alternatives. Where an attribute performs well relative to other alternatives, an improvement generates only a small decrease in regret; whereas the same magnitude of improvement generates a

larger decrease in regret if the attribute was performing relatively poorly to begin with. Consequently, the RRM model does not exhibit the property of independence from irrelevant alternatives, even with the assumption of *i.i.d.* error terms. Moreover, the RRM is just as parsimonious as the standard RUM model, unlike other models of contextual effects which typically require the estimation of additional parameters (Chorus, 2010). Despite these desirable properties, empirical support for the RRM, however, appears mixed. The RRM model only marginally outperforms its RUM counterpart in three out of the four datasets reported by Chorus (2010). In the remaining dataset, the linear additive RUM turns out to be the better model instead.

Embedding the compromise effect into discrete choice models is also possible through the proposed specifications offered by Kivetz *et al.* (2004). Their contextual models incorporate the use of reference points and also account for either loss aversion or concavity in gains in the context of the current choice set. Instead of a reference alternative, reference attribute levels are used. In their loss aversion model, the reference point is taken to be the mid-point of the attribute range of the alternatives in the local choice set, which is not necessarily equal to the attribute levels in the existing status-quo choice option. The value function is defined in Equation (7) as:

$$
\begin{aligned}
V_j = &\sum_k \Big[v_{jk}(X_{jk}) - v_{rk}(X_{rk}) \Big] \times 1\Big(v_{jk}(X_{jk}) \geq v_{rk}(X_{rk}) \Big) \\
&+ \sum_k \lambda_k \Big[v_{jk}(X_{jk}) - v_{rk}(X_{rk}) \Big] \times 1\Big(v_{jk}(X_{jk}) < v_{rk}(X_{rk}) \Big)
\end{aligned}
\tag{7}
$$

In Equation (7), V_j is the value of alternative j (given a choice set s), $v_{jk}(X_{jk})$ is the utility of attribute k of alternative j, λ_k is the loss aversion parameter for attribute k and X_{rk} indicates the reference value of attribute k in choice set s. If respondents display loss aversive tendencies, λ_k will be greater than one.

On the other hand, Kivetz *et al.*'s (2004) context concavity model takes the attribute value with the lowest part-utility as the reference point and codes the utility of other attribute values as gains against the reference. This model specification is shown in Equation (8):

$$
V_j = \sum_k \Big(v_{jk}(X_{jk}) - v_{rk}(X_{rk}) \Big)^{c_k}
\tag{8}
$$

As its name suggests, the context concavity model assumes that the utility gains are concave relative to the reference, as predicted by prospect theory. Hence, c_k is introduced as a concavity parameter for attribute k. X_{rk} in this case is the attribute value that gives the lowest utility on attribute k across all alternatives in the choice set. More generally, the concavity parameter implies diminishing marginal sensitivity to gains; thus, the in-between alternative with its moderate gains on the attributes benefits more compared to the extreme alternatives.

10

Another model, attributable first to Tversky and Simonson (1993) as a componential context model, and later updated by Kivetz *et al.* (2004) as a relative advantage model, is another candidate for explaining the compromise effect. This model is shown in Equation (9):

$$V_j = \sum_k v_k(X_{jk}) + \theta \sum_{j' \in S} R(j, j')$$

(9)

$R(j, j')$ denotes the relative advantage of alternative *j* over alternative *j'*, and ϑ is the weight given to the relative advantage component of the model. Tversky and Simonson (1993) define $R(j, j')$ as follows: First, for a pair of alternatives *(j, j')*, consider the advantage of *j* over *j'* with respect to an attribute *k*, denoted in Equation (10) by:

$$A_k(j, j') = \begin{cases} v_k(X_{jk}) - v_k(X_{j'k}) \text{ if } v_k(X_{jk}) \geq v_k(X_{j'k}), \\ 0 \text{ otherwise} \end{cases}$$

(10)

Define the disadvantage of *j* over *j'* with respect to an attribute *k* as an increasing convex function $\delta_k(.)$ of the corresponding advantage function $A_k(j', j)$, i.e., $D_k(j, j') = \delta_k(A_k(j', j))$. $\delta_k(.)$ is assumed to be convex due to loss aversion in evaluating a disadvantage. A functional form for $D_k(j, j')$, due to Kivetz *et al.* (2004), is suggested in Equation (11):

$$D_k(j, j') = A_k(j', j) + L_k A_k(j', j)^{\psi_k}$$

(11)

Here, L_k is a loss aversion parameter (*a priori* expected to be greater than zero) and ψ_k is a power parameter (*a priori* expected to be greater than one). The relative advantage of *j* over *j'* is then defined in Equation (12) as:

$$R(j, j') = \frac{\sum_k A_k(j, j')}{\sum_k A_k(j, j') + \sum_k D_k(j, j')} = \frac{A(j, j')}{A(j, j') + D(j, j')}$$

(12)

$R(j, j') = 0$ if choice set *s* contains two or less elements. Like the RRM model, the relative advantage model assumes that each alternative is compared against all other alternatives in the choice set.

Testing the various model forms for the compromise effect, Kivetz *et al.* (2004) find that the parameter estimates for all models are consistent with *a priori* expectations from theory. However, the contextual concavity model and the loss aversion model have superior measures of fit and predictive validity compared to the relative advantage model, with the latter performing no better than the standard model in a handful of cases. The poorer performance of the relative advantage

model may be a consequence of the way the model was estimated, as Kivetz *et al.* report difficulties with identifying the power parameters ψ_k of the highly non-linear disadvantage function.

Conceptually, however, the relative advantage model serves a useful purpose in highlighting a distinction between what Simonson (2008) calls "inherent" and "constructed" preferences. Simonson argues that there may be a more important role for stable or inherent preference values to play in shaping decisions than has been recognised in the behavioural decision literature. Equation (9) thus represents the notion that preferences are separable into inherent and constructed components (Kivetz *et al.*, 2008). The first term on the right hand side of Equation (9) represents value maximisation independent of contextual effects while the second term is context dependent utility. The parameter ϑ therefore represents the weight given to constructed preference via the choice context vis-à-vis inherent preferences. Furthermore, the parameter ϑ may be expressed as a function of respondent and choice task characteristics, allowing the weight of constructed preferences to vary across respondents and choice tasks.

Rooderkerk *et al.*'s (2011) model essentially follows this line of thinking by decomposing utility into additively separable context-free and context dependent components. Their key innovation is to consider the context dependent component of utility as a linear combination of three contextual effects that have been well-documented: the compromise effect, the attraction effect (Huber *et al.*, 1982) and the similarity effect (Tversky, 1972). Metrics for these effects are constructed using measures of Euclidean distance between attributes. They find that a model that controls for these contextual effects, compared to a RUM model without contextual effects, shows improvement in most of the fit statistics analysed.

Heuristics of Choice Set Inter-Dependence

The notion of "relational" can be extended to allow preceding choice tasks or choice outcomes to impact current choice. As noted by Simonson and Tversky (1992), "in deciding whether or not to select a particular option, people commonly compare it to other alternatives that are currently available as well as with relevant alternatives that have been encountered in the past" (Simonson and Tversky, 1992, p. 282). As most choice experiments require respondents to answer a series of choice tasks, this assertion implies that preferences over attributes are not necessarily independent across choice sets.

Indeed, we now have a significant amount of evidence that what respondents encounter in previous choice sets matters in the current decision making context. "Ordering anomalies", where choice is biased by the sequence of attribute values observed in the preceding choice set(s), are not uncommon (Day and Prades, 2010). For example, if a price attribute of one alternative is seen to increase from one choice set to another, the proportion of respondents choosing that alternative in the second choice set is smaller than if the choice sets were reversed. A proposed explanation is a 'good deal / bad deal' heuristic (Bateman *et al.*, 2008) whereby 'good deals' in the current choice task, relative to the price-attribute combinations encountered in previous choice tasks, are chosen

12

more frequently than relative 'bad deals'. A trade-off contrast (Simonson and Tversky, 1992) by which current preferences are revised on the basis of previous price or cost attributes may also explain ordering effects. This is a specific example of a more general phenomenon of preference reversal (Tversky *et al*., 1990).

Strategic misrepresentation has also been invoked as one justification for incorporating the attributes of some previously chosen non status-quo alternative as a reference point in the current choice set. The argument from a public goods provision context is that people aim to increase the likelihood of their most preferred alternative being implemented by deliberately withholding the truth about their preferences in the current choice task if chosen alternatives in previous choice tasks have better attribute values (such as lower cost) than those in the current choice task. Strategic misrepresentation assumes that the respondents have stable and well formed preferences, but that a discrepancy exists between stated preferences and underlying true preferences. A weaker version of strategic misrepresentation allows respondents to consider the likelihood that the good would not be provided if they do not reveal their true preferences, and hence to only reject truth-telling probabilistically (McNair *et al.*, 2011b). Other papers have also concluded that the data are consistent with respondents using previously encountered information and past choices and that estimated parameters are sensitive to how choice sets are ordered. Strategic behaviour in respondents, particularly the weaker version of strategic behaviour, is not rejected (Scheufele and Bennett 2010a; 2010b; McNair *et al.*, 2011a).

Another explanation for considering features of previously encountered choice sets into the current choice set involves a value learning heuristic, which assumes truth telling, but poorly formed initial preferences. Value learning involves the discovery of preferences and taste parameters may change according to the attribute levels presented to the respondent. Hence, preferences can be influenced by the starting point and subsequent attribute values (McNair *et al.*, 2011a), with the 'good-deal / bad-deal' heuristic being a specific case in point. It is possible that in a group of respondents, decision process heterogeneity involving several heuristics are at work and no one heuristic dominates. McNair *et al.* (2011a) show that responses to a sequence of binary choice tasks involving the provision of an underground electricity network are consistent with both a weak form of strategic misrepresentation and with a 'good deal/bad deal' heuristic, while in an equality constrained model of probabilistic decision processes, strategic misrepresentation and value learning can be modelled as distinct classes of heuristics for sub-groups of respondents, with the class membership probabilities of these heuristics estimated to be higher than the class membership probability for the utility function with the standard assumptions (McNair *et al.*, 2011b).

Underpinning value learning and strategic misrepresentation is the notion of reference point revision (DeShazo, 2002). In experiments which include the status quo as one of the alternatives, the oft-observed status quo bias (Samuelson and Zeckhauser, 1988; Fernandez and Rodrik, 1991) may mean that the status quo itself simply ends as the reference point. Hensher and Collins (2011) test whether reference points are shifted when non status quo alternatives are chosen and find that if a non-reference (i.e., non status quo) alternative is chosen in the preceding choice set $s-1$, the reference in the current choice set *s* is revised and the utility of the non status quo alternatives increases. This suggests a shift in the value function around a new reference point.

Beyond the environmental and transport applications, the marketing literature has also sought to understand how past experience with the attributes of a certain product affects current decision making. For example, Briesch *et al.* (1997) argue that when previously encountered attributes or alternatives are used as a reference, judgements are assumed to be memory-based because information is retrieved from memory and then compared to what is currently available in the choice set. Thus, a vastly superior dominant alternative encountered in a choice set creates conditions for memory-based judgments to take place and it seems likely that such an alternative will be held in memory as a reference point in future choice sets. Concerning high frequency purchases of consumer goods, Briesch *et al.* evaluate various econometric specifications of references involving past or current prices and find that a reference specification dependent on memory-based brand specific past prices provides the best overall model fit.

With a sufficient number of choice sets per respondent, such as when decisions of a panel of respondents are recorded over a period of time, it may be possible to test whether intervening choice sets matter. Heckman (1981) suggests the idea of modelling the dynamics of past influences on current decisions through a general model of structural state dependence and habit persistence. Structural state dependence occurs when previous choice outcomes affect current utility, whereas habit persistence allows for previous utility evaluations to affect current utility. Heckman's model is picked up by Swait *et al.* (2004). In the model specification for habit persistence, the current utility at time or choice set *s* of alternative *j* is defined through a meta-utility function as shown in Equation (13):

$$\hat{V}_{js} = \prod_{i=0}^{i=s} \alpha_{j,s-i} \exp(V_{j,s-i}) \tag{13}$$

Meta-utility $\hat{V}_{js}$ is dependent on all past (static) utilities $V_{j,s-i}$ which is itself dependent only on the attributes in the period $s - i$. The link between current utility and historical observed utilities is achieved through a path-dependence parameter $\alpha_{j,s-i}$, where $\alpha_{j,s-i}$ might also be interpreted as the weights associated with the previous periods. $\alpha_{j,s-i}$ satisfies the conditions $0 \le \alpha_{j,s-i} \le 1$; $\alpha_{js} = 1$. Taking logs to obtain a linear additive form and adding past and contemporaneous error terms results in Equation (14):

$$\ln(\hat{V}_{js}) = \sum_{i=0}^{i=s} V_{j,s-i} + \sum_{i=0}^{i=s} \ln(\alpha_{j,s-i}) + \sum_{i=0}^{i=s} \varepsilon_{j,s-i} \tag{14}$$

As the first right hand side term of Equation (14) contains all past attribute levels, this equation can also be seen to link "current utility to historical observed attribute levels in a fashion that is consistent with learning about attributes or updating" (Swait *et al.*, 2004, p. 98). Attribute levels in

previous periods are therefore combined with current attribute levels in a form of temporal averaging. Incorporating state dependence into the model involves using a dummy variable that equals 1 for alternative *j* in current choice set *s* if the same alternative had been chosen in choice set *s − 1*. The variance structure of the disturbance term can be allowed to vary over time, providing a form of temporal heteroscedasticity. With single respondents answering repeated choice experiments, this model provides another way of investigating the role of the value learning heuristic.

Heckman's idea of state dependence implies that the utility for alternatives in the current state or choice set is directly modified by previous choices or previously encountered attributes. One difficulty with this approach is the potential endogeneity that might arise as a result of the error term being correlated with some of the co-variates. The use of lagged endogenous variables to model habits is avoided in Adamowicz and Swait (2011) through a two stage decision process. In their model, the higher level decision involves a choice from one of various decision strategies. These decision strategies correspond to: (i) choosing any alternative that is different from the last alternative chosen (variety seeking heuristic); (ii) choosing the same alternative again (habit heuristic); or (iii) full evaluation (utility maximisation in a fully compensatory sense).The second stage of the decision process is the evaluation of alternatives conditioned on the decision strategy of the first stage. Past choice behaviour influences the higher level decision on which decision strategy to adopt, but does not directly affect the evaluation of the alternatives themselves. When applied to a dataset of routine and repeated purchases of various consumer goods over time, Adamowicz and Swait find that their proposed model is preferred over a state dependence model on the basis of information criteria and on the plausibility of direct price elasticities. By not correcting for habit persistence where it is more likely, the standard state dependence model leads to an over-prediction of price elasticity.

Another way of linking the attribute levels in the preceding choice set to the attributes in the current choice set involves a just noticeable difference heuristic (Cantillo *et al.*, 2006). A change in the attribute level from choice task $s-1$ to choice task *s* is assumed to be perceptible to the respondent if the magnitude of the change in the attribute levels of attribute *k* exceeds a certain threshold, i.e., $\left|\Delta X_{k,s}\right| = \left|X_{k,s} - X_{k,s-1}\right| \geq \delta_k$, for non-negative threshold values δ_k of attribute *k*. Like several of the threshold formulations described earlier, thresholds can be assumed to be individual-specific, randomly distributed across the population, and may also depend on socio-demographic characteristics.

Cantillo *et al.* assume that respondents only perceive the part of the attribute level change that is bigger than the threshold, as in Equation (15) below:

$$X_{jks} = X_{jk,s-1} + sgn(\Delta X_{jks})\max(\left|\Delta X_{jks}\right| - \delta_k, 0) \tag{15}$$

If *m* out of the K attributes are associated with a perception threshold, utility can be written in Equation (16) as:

15

$$U_{js} = V_{js} + \varepsilon_{js} = \sum_{k=1}^{m} \beta_{jk} X_{jks} + \sum_{k=m+1}^{K} \beta_{jk} X_{jks} + \varepsilon_{js}$$

$$= \sum_{k=1}^{m} \beta_{jk} \left[X_{jk,s-1} + \Delta X_{jks} \left(1 - \frac{\delta_k}{\left| \Delta X_{jks} \right|} \right) I_{jk} \right] + \sum_{k=m+1}^{K} \beta_{jk} X_{jks} + \varepsilon_{js}$$

(16)

where $I_{jk} = \begin{cases} 1 \text{ if } \left| \Delta X_{jks} \right| \geq \delta_k \\ 0 \text{ otherwise} \end{cases}$

To complete the model, a joint density function needs to be assumed for δ_k. The value for *m* is then determined exogenously, by testing each attribute against a threshold constrained assumption until the best model fit is obtained. Applying this model to a route-choice stated preference survey for car trips, Cantillo *et al.* conclude that a threshold exists for the travel time attribute but not for the cost or variability attribute. Assuming a threshold model results in parameter estimates that imply a lower value of travel time savings, on average, compared to a traditional MNL model.

Cantillo *et al.*'s just noticeable difference heuristic provides one way of allowing respondents to modify the attribute values presented to them in a particular choice task, thereby relaxing the frequently maintained assumption in most choice models that respondents take the attribute levels as given. In applications where variability matters, for example, in transport where both travel times and variability of travel times are important determinants of choice (see Hensher and Li, 2010), the travel time attribute may itself be changed or edited by the respondent, with the magnitude of the edit possibly depending on the variability attribute and any associated threshold.

Other Approaches

Heterogeneity in decision processes may be inferred from observed choice outcomes by directly embedding various heuristics into the modelled component of utility functions in latent class or probabilistic decision process models (Hensher and Collins, 2011; Hess *et al.*, 2011; Hensher and Greene, 2010; McNair *et al.*, 2011b; Scarpa *et al.*, 2009). In transport applications, these models have been used to test the heuristics of common-metric attribute aggregation, attribute non-attendance and decision rules like majority of confirming dimensions to explain choice in the context of a toll road/non toll road alternative (Hensher, 2010; Hensher and Collins, 2011). The lexicographic and EBA decision rules may also be modelled using this approach (see for example, Hess *et al.*, 2011). The overall conclusion from the research is that accounting for decision process heterogeneity in latent classes – that is, allowing heuristic use to vary by subgroups of respondents (up to a probability) – leads to improvements in model fits compared to the standard multinomial logit model, and where assessed, the mixed logit model based on a standard utility specification. In instances where supplementary questions are a part of the survey instrument, directly embedding self-stated responses to questions of whether certain attributes are ignored has further improved the explanatory power of the models and the efficiency of marginal WTP estimates (see for example, Scarpa *et al.*, 2010).

Hierarchical Bayes modelling has also proven useful when estimating heterogeneity in cases where the decision sequence, constraints and thresholds are latent. Gilbride and Allenby (2004) model the two stage decision processing strategy by assuming that screening rules exist to restrict a larger choice set into a smaller subset of alternatives for final evaluation. The screening rules considered include (i) a compensatory screening rule, where the deterministic portion of utility in the traditional compensatory sense must exceed a threshold; (ii) a conjunctive screening rule, where all attribute values must be acceptable and (iii) a disjunctive rule, where at least one attribute value needs to be acceptable. These thresholds are determined endogenously and are allowed to vary by respondents, not unlike the approach taken by Swait (2009) and Cantillo and Ortuzar (2005). In conclusion, Gilbride and Allenby (2004) find that the conjunctive screening model explains the data best. Further extensions of this work include modelling the EBA processing rule and an economic screening rule which stops all processing of an alternative whenever an undesirable attribute level is present (Gilbride and Allenby, 2006).

Conclusion

This review demonstrates that an active amount of research has been undertaken in the discrete choice modelling literature towards modelling the contribution of decision heuristics and contextual effects in explaining choice behaviour. The empirical evidence consistently shows that embedding heuristic rules into the modelled component of utility leads to improved measures of fit. However, given the plethora of heuristics that has been shown to exist, it may now be an opportune time to consider testing how multiple heuristics may be embedded in choice models. Modelling certain aspects of heterogeneity in decision rules might well complement the current research focus of modelling preference heterogeneity through more advanced discrete choice models such as the mixed multinomial logit.

As a practical way forward, future work might consider the use of mixture models, where multiple heuristics are weighted in a utility function, using weighting functions that depend on the socio-economic characteristics of the respondent and other choice context variables, including individual specific perceptions data, where available. These mixture models thus allow a certain degree of heterogeneity in the use of decision rules. Such an approach is especially relevant in testing cases such as the relative advantage model, where respondents' utility for a certain alternative are assumed to be a function of both inherent preference (predispositions) and constructed preferences which depend on the choice context.

A related observation is that in contrast to prospect theory, there appears to be comparatively much less research done on the application of models of extremeness aversion (the compromise effect) and other choice set contextual components within the discrete choice literature in transportation. This lack of research is all the more remarkable considering that authors in the management science and marketing fields have consistently identified these effects as important determinants of choice. Fortunately, testing these heuristics is highly feasible even with existing datasets. Compared to models of thresholds, no additional information on thresholds needs to be collected from respondents, nor do rules on thresholds need to be assumed and tested.

In addition, much research in this area involves using the same dataset to estimate various models with decision heuristics embedded and comparing these to a standard model which is usually specified to be MNL, fully compensatory, linear-in-the-parameters and linear-in-the-attributes. There is comparatively less work done on testing the independence of heuristics across multiple datasets, in other words, testing if the heuristics that purportedly explain the data better in one dataset can do so likewise for other datasets. This line of inquiry would be especially pertinent for models which rely on additional data for identification (such as models using self-reported threshold data). It will be useful to have a better measure of how general is the use of such heuristics, before starting on a data collection process that might be resource intensive.

The heuristics that have been reviewed in this paper have been developed and assessed within specific discrete choice modelling frameworks; however while the evidence is limited to static models, it does signal a need to consider incorporating the findings into other behavioural frameworks that incorporate choice model components such as activity- based models and even model systems that allow for dynamic factors (learning, adaptation, habitual behaviour), and mental processes that have an important role in heuristic decision making. Wherever a utility expression is included in a model system, the evidence herein suggests that value might be added if new or alternative functional forms that capture a range of decision making rules or heuristics are specified.

In a similar vein, the literature has observed that model outputs such as welfare estimates and marginal willingness to pay can be substantially different when the model departs from the standard assumptions about decision making. However, there is mixed evidence as to the direction of change to willingness to pay measures when heuristics are embedded into choice models. For example, some papers suggest an increase in the value of travel time savings when heuristics are modelled (e.g., Hensher, 2010), while others come to the opposite conclusion (e.g., Cantillo and Ortuzar, 2005). This is another area that warrants more research attention.

References

Abou-Zeid, M. and Ben-Akiva, M. (2011) The effect of social comparisons on commute well-being, *Transportation Research Part A*, 45(4), pp. 345-361.

Adamowicz, W. and Swait, J. (2011) Are food choices really habitual? Integrating habits, variety seeking and compensatory choice in a utility maximising framework, paper presented at the Second International Choice Modelling Conference, Leeds, UK.

Andersen, S., Harrison, G.W., Lau, M.I. and Rutström, E.E. (2007) Dual criteria decisions, *Working Paper 06-11*, Department of Economics, College of Business Administration, University of Central Florida.

Avineri, E. and Prashker, J. N. (2004) Violations of expected utility theory in route-choice stated preferences: certainty effect and inflation of small probabilities, *Transportation Research Record*, 1894, pp. 222-229.

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

Bell, D.E. (1982) Regret in decision making under uncertainty, *Operations Research*, 30(5): pp. 961-981.

Bettman, J. R., Luce, M.F., and Payne, J. W. (1998) Constructive consumer choice processes. *Journal of Consumer Research*, 25(3), pp. 187-217.

Briesch, R.A., Krishnamurthi, L., Mazumdar, T., and Raj, S.P. (1997) A comparative analysis of reference price models, *Journal of Consumer Research*, 24(2), pp. 202-214.

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

Cantillo, V. and Ortuzar, J.de D. (2005) A semi-compensatory discrete choice model with explicit attribute thresholds of perception, *Transportation Research Part B*, 39(7), pp. 641-657.

Chorus, C.G. (2010) A new model of random regret minimization, *European Journal of Transportation and Infrastructure Research*, 10(2), pp. 181-196.

Chorus, C.G., Arentze, T.A. and Timmermans, H.J.P. (2008) A random regret minimization model of travel choice, *Transportation Research Part B*, 42(1), pp. 1-18.

Day, B. and Prades, J.P. (2010) Ordering anomalies in choice experiments, *Journal of Environmental Economics and Management*, 59(3), pp. 271-285.

DeShazo, J.R. (2002) Designing transactions without framing effects in iterative question formats, *Journal of Environmental Economics and Management*, 43(3), pp. 360-385.

Fernandez, R. and Rodrik, D. (1991) Resistance to reform: status quo bias in the presence of individual-specific uncertainty, *American Economic Review*, 81(5), pp. 1146-1155.

Gigerenzer, G., Todd, P.M. and the ABC Research Group (1999) *Simple Heuristics that Make Us Smart* (New York, NY: Oxford University Press).

Gilbride, T.J. and Allenby, G.M. (2004) A choice model with conjunctive, disjunctive and compensatory screening rules, *Marketing Science*, 23(3), pp. 391-406.

Gilbride, T.J. and Allenby, G.M. (2006) Estimating heterogeneous EBA and economic screening rule choice models, *Marketing Science*, 25(5), pp. 494-509.

Greene, W.H. and Hensher, D.A. (2012) Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model, *Applied Economics*, doi:10.1080/00036846.2011.650325.

Heckman, J. (1981) Statistical models for discrete panel data, in C. F. Manski and D. McFadden (Eds.) *Structural Analysis of Discrete Data and Econometric Applications*, pp. 114-177 (Cambridge, MA: The MIT Press).

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

Hensher, D.A. (2010) Attribute processing, heuristics and preference construction in choice analysis, in S. Hess and A. Daly (Eds.) *Choice Modelling: The State-of-the-Art and the State-of-Practice – Proceedings from the Inaugural International Choice Modelling Conference*, pp. 35-70 (Bingley, UK: Emerald Group Publishing).

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

Hensher, D.A. and Greene, W.H. (2010) Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification, *Empirical Economics*, 39(2), pp. 413-426.

Hensher, D.A. and Li, Z. (2010) Valuing travel time variability within a rank-dependent utility framework and an investigation of unobserved taste heterogeneity, *Journal of Transport Economics and Policy*, available online at http://www.ingentaconnect.com/ content/lse/jtep/pre-prints/hensherandliversion2.

Hensher, D.A. and Rose, J.M. (2011) The influence of alternative acceptability, attribute thresholds and choice response certainty on automobile purchase preferences, *Journal of Transport Economics and Policy*, available online at http://www.ingentaconnect.com/content/ lse/ jtep/pre-prints/hensherandrose

Hess, S., Stathapoulos, A. and Daly, A. (2011) Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies, *Transportation*, doi:10.1007/s11116-011-9365-6

Hole, A.R. (2011) A discrete choice model with endogenous attribute attendance, *Economics Letters*, 110(3), pp. 203-205.

Huber, J., Payne, J.W. and Puto, C. (1982) Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis, *Journal of Consumer Research*, 9(1), pp. 90-98.

Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decision under risk, *Econometrica*, 47(2), pp. 263-292.

Kivetz, R., Netzer, O. and Schrift, O. (2008) The synthesis of preference: bridging behavioral decision research and marketing science, *Journal of Consumer Psychology*, 18(3), pp. 179-186.

Kivetz, R., Netzer, O. and Srinivasan, V. (2004) Alternative models for capturing the compromise effect, *Journal of Marketing Research*, 41(3), pp. 237-257.

Li, Z. and Hensher, D.A. (2011) Prospect theoretic contributions in understanding traveller behaviour under risk or uncertainty: a review and some comments, *Transport Reviews,* 31(1), pp. 97-117.

Loomes, G. and Sugden, R. (1982) Regret theory: an alternative theory of rational choice under uncertainty, *The Economic Journal*, 92(368), pp. 805-824.

Louviere, J.J. and Meyer, R.J. (2008) Formal choice models of informal choices – what choice modelling research can (and can't) learn from behavioural theory, in N. K. Malhotra (Ed.) *Review of Marketing Research Volume 4*, pp. 3-32 (Armonk, NY: M.E. Sharpe).

Manski, C.F. (1977) The structure of random utility models, *Theory and Decision*, 8(3), pp. 229-254.

McFadden, D. (1999) Rationality for economists? *Journal of Risk and Uncertainty*, 19(1-3), pp. 73-105.

McNair, B.J., Bennett, J. and Hensher, D.A. (2011a) A comparison of responses to single and repeated discrete choice questions, *Resource and Energy Economics*, 33(3), pp. 554-571.

McNair, B.J., Hensher, D.A. and Bennett, J. (2011b) Modelling heterogeneity in response behaviour towards a sequence of discrete choice questions: a probabilistic decision process model, *Environment and Resource Economics*, doi: 10.1007/s10640-011-9514-6.

Payne, J.W. and Bettman, J.R. (1992) Behavioural decision research: a constructive processing perspective, *Annual Review of Psychology*, 43, pp. 87-131.

Payne, J.W., Bettman, J.R. and Johnson, E.J. (1993) *The Adaptive Decision Maker* (Cambridge, UK: Cambridge University Press).

Payne, J.W., Bettman, J.R. and Schkade, D.A. (1999) Measuring constructed preferences: towards a building code, *Journal of Risk and Uncertainty*, 19(1-3), pp. 243-270.

Rooderkerk, R.P., Van Heerde, H.J. and Bijmolt, T.H.A. (2010) Incorporating context effects into a choice model, *Journal of Marketing Research*, 48(4), pp. 767-780.

Russo, J.E. and Dosher, B.A. (1983) Strategies for multiattribute binary choice, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 9(4), pp. 676-696.

Samuelson, W. and Zeckhauser, R. (1988) Status quo bias in decision making, *Journal of Risk and Uncertainty*, 1(1), pp. 7-59.

Scarpa, R., Gilbride, T.J., Campbell, D. and Hensher D.A. (2009) Modelling attribute non-attendance in choice experiments for rural landscape valuation, *European Review of Agricultural Economics*, 36(2), pp. 151-174.

Scarpa, R., Thiene, M. and Hensher, D.A. (2010) Monitoring choice task attribute attendance in nonmarket valuation of multiple park management services: does it matter? *Land Economics*, 86(4), pp. 817-839.

Scheufele, G. and Bennett, J. (2010a) Effects of alternative elicitation formats in discrete choice experiments, *54th Annual Australian Agricultural and Resource Economics Society Conference*, Adelaide, Australia, available online at http://purl.umn.edu/59158.

Scheufele, G. and Bennett, J. (2010b) Ordering effects and response strategies in discrete choice experiments, *2010 World Congress of Environmental and Resource Economists*, Montreal, Canada, available online at http://purl.umn.edu/107743.

Senbil, M. and Kitamura, R. (2004) Reference points in commuter departure time choice: a prospect theoretic test of alternative decision frames, *Journal of Intelligent Transportation Systems*, 8(1), pp. 19-31.

Simon, H. A. (1955) A behavioral model of rational choice, *Quarterly Journal of Economics*, 69(1), pp. 99-118.

Simonson, I. (2008) Will I like a "medium" pillow? Another look at constructed and inherent preferences, *Journal of Consumer Psychology*, 18(3), pp. 159-169.

Simonson, I. and Tversky, A. (1992) Choice in context: tradeoff contrast and extremeness aversion*, Journal of Marketing Research*, 29(3), pp. 281-295.

Stevenson, M.K., Rusemeyer, J.R. and Naylor, J.C. (1990) Judgment and decision-making theory, in M.D. Dunnette and L.M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology*, *Vol. I*, pp. 283-374 (Palo Alto, CA: Consulting Psychologists Press).

Suzuki, Y. (2007) Modeling and testing the "two-step" decision process of travelers in airport and airline choices, *Transportation Research Part E*, 43(1), pp. 1-20.

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

Swait, J. (2009) Choice models based on mixed discrete/continuous pdfs, *Transportation Research Part B*, 43(7), pp. 766-783.

Swait, J. and Adamowicz, W. (2001) 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*, 86(2), pp. 141-167.

Swait, J., Adamowicz, W. and van Buren, M. (2004) Choice and temporal welfare impacts: incorporating history into discrete choice models, *Journal of Environmental Economics and Management*, 47(1), pp. 94-116.

Swait, J. and Ben-Akiva, M. (1987) Incorporating random constraints in discrete models of choice set generation, *Transportation Research Part B*, 21(2), pp. 91-102.

Timmermans, H. (2010) On the (ir)relevance of prospect theory in modelling uncertainty in travel decisions, *European Journal of Transport and Infrastructure Research*, 10(4), pp. 368-384.

Tversky, A. (1969) Intransitivity of preferences, *Psychological Review*, 76(1), pp. 31-48.

Tversky, A. (1972) Elimination by aspects: a theory of choice, *Psychological Review*, 79(4), pp. 281-299.

Tversky, A. and Kahneman, D. (1991) Loss aversion in riskless choice: a reference-dependent model, *Quarterly Journal of Economics*, 106(4), pp. 1039-1061.

Tversky, A., Slovic, P. and Kahneman, D. (1990) The causes of preference reversal, *American Economic Review*, 80(1), pp. 204-217.

Tversky, A. and Simonson, I. (1993) Context-dependent preferences, *Management Science*, 39(10), pp. 1179-1189.

Van de Kaa, E.J. (2010a) Applicability of an extended prospect theory to travel behaviour research: a meta-analysis, *Transport Reviews*, 30(6), pp. 771-804.

Van de Kaa, E.J. (2010b) Prospect theory and choice behaviour strategies: review and synthesis of concepts from social and transport sciences, *European Journal of Transport and Infrastructure Research*, 10(4), pp. 299-329.

Young, W. (1984) A non-tradeoff decision making model of residential location choice, *Transportation Research Part A*, 18A(1), pp. 1-11.

Zhang, J. and Adamowicz, W.L. (2011) Unraveling the choice format effect: a context-dependent random utility model, *Land Economics*, 87(4), pp. 730-743.