I T L S

WORKING PAPER

ITLS-WP-07-22

Selective developments in choice analysis and a reminder about the dimensionality of behavioural analysis

By

David A Hensher, John M Rose & Sean Puckett

November 2007

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-07-22

TITLE: **Selective developments in choice analysis and a reminder about the dimensionality of behavioural analysis**

ABSTRACT: Developments in data and modeling paradigms in choice analysis are occurring at a fast pace. A review of activity leading up to each IATBR conference shows progress on many fronts. This paper takes a selective view of some of these developments, especially those that have been close to the research program of the authors. We focus on four broad themes – information processing strategies, especially in the context of stated choice studies; agency interdependency (with a strong applied focus), developments in the design of choice experiments, and a smorgasbord of themes centered on expanding the behavioral capabilities (and longer term forecasting accuracy) of discrete choice models, especially in terms of their recognition of ways of accommodating the other themes in the paper.

KEY WORDS: ***Choice, attribute processing, efficient designs, joint decision making, willingness to pay***

AUTHORS: David A Hensher, John M Rose & Sean Puckett

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: November 2007

1. Introduction

Choice analysis has become mainstream in transportation research. Beginning with a relatively narrow focus on the development of econometric models to estimate the parameters of discrete choice models, the literature has evolved into a number of streams, all having real potential to be integrated in a *behavioral system of choice and payoff*.

While noting the fast pace of development of model specification, from simple (but still useful) multinomial logit to ever-increasingly detailed closed and open form choice models, that account for the sources of observed and unobserved random and systematic heterogeneity and heteroskedasticity in preferences, we are now seeing a growing interest in the choice *process* in complementing the dominant emphasis on the choice *outcome*[1]. Indeed some researchers with a penchant beyond mainstream economics (notably in behavioral economic psychology, and to a lesser extent human geography and sociology), have been arguing for many years that the economist's perspective on choice and random utility is limiting; not so much because it is wrong, but more that it imposes bounds that are somewhat narrow in what can be incorporated in the study and modeling of choice. Choice after all has no disciplinary bounds – it is strictly behavioral.

Themes that have emerged in recent years have highlighted the need to refocus the boundaries and to give as much credence to broadened themes, as we continue to deliver in extending the econometric niceties of the family of choice models, especially the set of logit derivatives. This paper synthesizes a number of themes that are exercising the minds of a growing number of travel behavior researchers. In one sense, these themes are far from new, but the research effort is growing significantly, and we see very strong signs that the new insights can be relatively easily integrated into a system of choice models that recognize not only outcome but also process. The intent clearly is to justify the added 'complexity' in terms of an improved understanding of decision making, and especially in improving our ability to predict behavioral response under conditions of change. Strictly, we promote 'relevancy' instead of 'complexity', since the latter has produced a belief in limiting empirical efforts for (unfounded) fear of cognitive burden (assumed to be highly correlated with the amount of information to report – especially in RP studies and process – especially in SP studies).

The themes presented in this review promote a greater focus on the way that information is processed in choice making, the link between the amount of information on offer (especially that associated with the attributes within a stated choice (SC) framework) and its relevancy, and whether it is ignored or rearranged for a variety of rational reasons. We also promote the importance of joint decision making and ways in which barriers to cooperation can be identified and acted upon. Two additional topics that interweave, with a capability to embed information processing strategies and agent interdependency in choice outcome determination, are the design of SC experiments (which is now an order of magnitude more sophisticated) and the paths in advanced discrete choice models that are available to capture the role of attribute processing and agency dependency.

[1] Design hopefully to improve forecasting accuracy through an improved understanding on choice making.

The paper is organized as follows. In section 2, we present a framework within which information processing can be captured, that has a logical interface with choice analysis. This is followed in section 3 with a way of identifying the power relationship between agents where a cooperative outcome is necessary (often through concession) in order to activate a choice outcome. Section 4 overviews developments in the design of choice experiments, moving beyond orthogonal designs to designs that permit behaviorally plausible correlation (to some extent); in recognition of the need to deliver asymptotically efficient parameter estimates and cost-justifiable sample sizes. Section 5 links the elements of information processing to the recent promotion of reference classes, and section 6 offers suggestions on continuing research directions.

2. Information processing

"What lies ahead for discrete choice analysis? ... The potentially important roles of information processing, perception formation and cognitive illusions are just beginning to be explored and behavioral and experimental economics are still in their adolescence." (McFadden 2001)

Decision making life can be thought of as starting as a set of continuous random variables; and if there is no information added, the outcomes or payoffs are strictly random events. Fortunately, decision making is assisted by a number of behavioral inputs, often called attributes, but more generally a suite of cues and a set of rules used by individuals to assist them in processing the information centered on the cues in arriving at outcomes that deliver payoffs. Crucially, the payoffs result from the amount of information processed (Berg 2005). The decision making environment can be defined as a joint probability distribution over states of nature (i.e., alternatives on offer) and cues and a payoff function[2] that ranks stochastic outcomes conditional on observed cues and actions. Cues are typically a set of attributes and actions are the mechanisms that individuals adopt in processing the attributes to arrive at outcomes that have payoffs.

What we have found in recent years is that these attributes are the centerpiece of information processing and they can be relevant or not relevant (Hensher in press a,b). Within the relevant set, they can be processed or ignored, and they can be ignored in the presence or absence of cognitive constraints. Likewise deeming attributes as not relevant can be associated with cognitive and non-cognitive constraints. Importantly we argue that relevancy and ignoring such information are not contradictions in a behavioral sense. We are of the view that ignoring can be good (e.g., the divided attention syndrome), it can be smart and it aligns with a sentiment that individuals adopt relatively frugal action rules. Herbert Simon in the 1950s made the equivalent case that arguments that tend (in our view too much) to rely on cognitive burden to justify simplistic SC experiments, have failed to understand the simple principle that cognitive processes should not be evaluated in a vacuum, and that a context is required to establish how adaptive choice rules are.

Fundamentally, information is relevant if it contributes in a non-marginal way (i.e., beyond the just noticeable difference threshold) to payoff and the benefits perceived to flow through from effort expended in accounting for that attribute exceed the costs. As a

[2] A payoff function refers to the contrast of all costs and benefits linked to information processing in the context of a set of assessable attributes and outcomes. One interpretation, adopted in choice analysis, is that it is a utility function representing a preference ordering over all alternatives in a pre-defined choice set.

corollary, relevant information can be ignored within the context of good choice making for many non-trivial reasons. Within the choice making context, we distinguish between alternatives in a choice set of outcomes (i.e., what is the common behavioral metric in choice analysis) and the choice set of actions or decision rules on information processing that an individual selects to maximize the expected payoff function conditional on observable information. Central to the selection of a preferred information processing rule (described as an optimal action in behavioral economics) is the treatment of attributes (which we refer to as the *Attribute Processing Strategy* (APS)). This is at the heart of the process model, regardless of whether the attributes are pre-specified as in the majority of SC experiments or whether they are elicited through other mechanisms.

Drawing on the experimental findings from psychology, we know that individuals often make incomplete use of available information, which implies that, although expected payoff functions may be influenced by specific attributes, an adopted information processing rule does not depend on these specific attributes. Such information processing rules are incomplete in the sense that human cognition provides filters[3] that result in adaptive responses to specific types of payoff/information environments. This is not the result of bounded rationality per se[4] but the interaction of such rationality with the payoff-probability structure within the choice environment under study. Hence, ignoring attributes is a rational outcome of a choice process.

Cognitive constraints are commonly cited as the generic basis of information processing rules. Such constraints are derivatives of complex phenomena, many of which are unknown and/or poorly understood or articulated by the individual and/or the researcher. There is a large body of research that associates cognitive constraints with memory limitations (Stroop 1935), bounds on processing speed, pre-attentive capability (i.e., number of channels), all leading to mechanisms to cope or economize on processing resources such as ignoring certain amounts of information. The information ignored (or what is increasingly referred to as information suppression – see Erber and Fiske 1984) can include total exclusion of a specific attribute (with the selection rule being systematic[5] or random[6]), or partial exclusion as a consequence of noting but discounting its presence. Importantly, and almost perceptually self-evident, individuals discount specific information in line with their own payoff-probability function, which processes all opportunities and assigns a probability to each, which is mapped into an expected payoff to arrive at a preferred outcome or choice. The idea of *ignoring* information is complex in that individuals often analyze all attributes (or cues) in order to identify which ones can be excluded from the optimal action plan[7]. Within the setting of a (subjective) utility function, as is the case of SC analysis, the two-stage process applies. However, if we can assume the existence of an adaptive mechanism with a

[3] This enables cognitive effort in general and hence selective cognitive responses, to be allocated to the important tasks.

[4] Bounded rationality in economics is typically given a narrow interpretation often linked to coping in a negative sense (or sub-optimal sense); whereas a more appealing interpretation credits it as an adaptive mechanism to support enhanced outcomes. We read rationality as the product of behaviour *and* reasoning.

[5] For example, the level does not differentiate enough from a reference alternative or accumulated experience on expected gain, as postulated in case-based decision theory which invokes similarity weights (Gilboa and Schmeidler 2001).

[6] For example, ignore any attributes below the first three listed.

[7] This plan becomes the equilibrium state (at least in the short run), which is often a habit-forming state that is repeated without any further filtering tasks. This accords also with transactions cost theory and search/minimum-regret theory in economics. Any major changes in the context may invoke a review of the equilibrium state, leading to a renewed filtering process.

history of evolutionary exposure and (overt) experience, then it is possible to 'go straight to the set' that is used in selecting the optimal action plan (i.e., choice outcome). Another way of stating this is that when the payoff function is defined strictly by adaptation, then the optimal outcome does not depend on a current attribute, and hence the need to know about attribute processing in the current state is of little consequence to the outcome. Unfortunately the analyst is unlikely to know this and be able to make inferences up to a probability, and will have to rely on an explicit test of information processing involving a mixture of adaptive presence and current state attribute relevance.

The analyst has the task of identifying the components of the process-outcome model that drive individual decision making and the diversity of such models, as a way of accommodating the heterogeneity existing within a population of decision makers. Within the travel behavior setting, we are embellishing the SC framework to accommodate such features of decision making.

2.1 The stated choice setting: in need of revision?

Stated choice experiments are typified by a pre-determined set of attributes and alternatives, with the number of levels and range of each attribute fixed within the design. The experimental design is then developed under a specific set of rules, such as orthogonality or D-efficiency (see section 4), with or without priors on the parameter estimates, and typically under the assumption that the resulting data will be estimated under the multinomial logit IID condition. Although there has been a growing recognition that the design of choice experiments should be conditioned by the specific functional specification of the estimation model, another strand of activity is focusing on the influence that the dimensionality of the choice experiment has on the revelation of preferences and hence on choice responses.

This section discusses ways that the information in a SC experiment is processed; which is attributed in part to the dimensionality of the SC experiment and in part to recognition that there is substantial heterogeneity in the processing strategies of individuals in a sample. In particular we argue that failure to take into account the *relevancy* of the information offered in the evaluation process leading to a choice outcome, no matter how 'simple' or 'complex' a design is, will contribute to biases in preference revelation. The great majority of researchers and practitioners ignore this aspect of SC methods, assuming that attributes offered are all relevant to some degree (Hensher in press-b).

In recent years there has been a growing interest on understanding the processes or rules invoked by respondents in dealing with the information in SC experiments. Although the impetus for this focus appears to have been motivated by an interest in cognitive burden, research by Hensher (amongst others) found that the real issue is not the amount of information to process, which became associated with 'complexity', but rather the *relevance* of the information. This opened up the possibility that a study of the implications on choice response of the amount of information provided in a choice experiment should be investigated in the context of the broader theme of what rules individuals bring to bear when assessing the information in a choice experiment. These rules may be embedded in prejudices that have little to do with the amount of information in the experiment; rather they may be rational coping strategies that are used in everyday decision making for a whole host of reasons. There is an extensive literature on information processing, which includes prospect theory (Kahnemann and

Tversky 1979), case-based decision theory (Gilboa and Schmeidler 2001) and non-expected utility theory (Starmer 2000).

2.2 How does a respondent assess a stated choice task?

Imagine that you have been asked to review the following choice screen and indicate which alternative is your preferred (Figure 1). There is a lot of information in this screen that you have to *attend to*, in deciding what influences your decision (what we refer as *relevant* information). There are likely to be many implicit and often subconscious rules being adopted to process the attributes and alternatives that are used, possibly to cope in a constructive way with the amount of information to assess (what we refer to as a *coping* strategy). The screen, for example, may be regarded as too *complex* in terms of the amount of information presented and its content. Whether one invokes a specific set of *processing rules* to cope with complexity, or whether these are a subset of the rules you have built up over time and draw on from past experiences, may be unclear. What we do suspect is that there are a large number of processing rules (what we call heterogeneity in information processing) being used throughout any sampled population, and that individuals are using them to handle mixtures of relevancy and cognitive burden (including task learning)[8]. Indeed it may be reasonable to suggest that relevancy is in part a natural response to cognitive constraint (as suggested above).

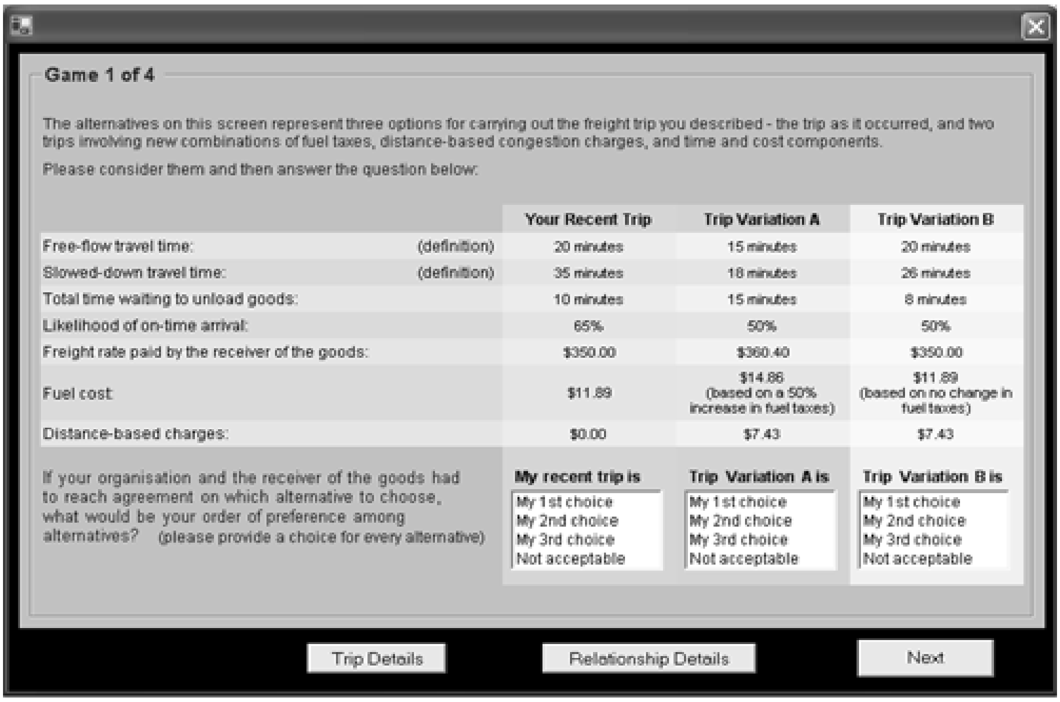


Figure 1: Example of a stated choice screen

It is reasonable to propose that individuals do have a variety of attribute processing (AP) styles, including the *simplifying* strategy of ignoring certain attributes (for whatever reason). Heterogeneity in AP strategies is widely reported in consumer research (see for

[8] Studies decomposing random parameters by amount of time spent, show (1) people spend longer on earlier choice sets and (2) the amount of processing time is a significant decomposition parameter for random parameter distributions.

example Hensher 2004; DeShazo and Fermo 2002, 2004) and its existence in choice experiments is supported by observation of lexicographic[9] choice behavior in segments of respondents completing SC surveys (see for example, Saelendsminde 2002)[10]. When researchers fail to account for such an AP strategy, they are essentially assuming that all designs are comprehensible, all design attributes are relevant (to some degree) and the design has accommodated the relevant amount of 'complexity' necessary to make the choice experiment meaningful (Hensher *et al.* 2005). Ideas of good and smart choosing appear absent.

Experimental evidence and self-reported decision protocols support the view that heuristic rules are the proximate drivers of most human behavior (McFadden 2001). The question remains as to whether rules themselves develop in patterns that are broadly consistent with random utility maximization postulates. If there are preferences behind rules, then it is possible to recover them and correctly evaluate policies in terms of these underlying preferences. If not, economics will have to seek a new foundation for this task. While many psychologists argue that behavior is far too sensitive to context and affect to be usefully related to stable preferences, this is a somewhat pessimistic view. A number of authors have challenged this position (e.g., Hensher in press, McFadden 2001, Swait and Adamowicz 2001). Many behavioral deviations from the economist's standard model can be attributed to perceptual illusions, particularly in the way that we process information, rather than a more fundamental breakdown in the pursuit of self-interest. Many of the rules we do use are essentially defensive, protecting us from mistakes that perceptual illusions may induce.

The (implicit) assumption in SC studies that all attributes are processed by all respondents has been challenged by a number of researchers (e.g., DeShazo and Fermo 2004, Hensher 2004, in press, Hensher *et al.* 2005) who argue that it is more likely that individuals react to increasingly 'complex' choice situations by adopting one of two AP strategies, broadly defined by the rival *passive bounded rationality* and *rationally-adaptive* behavioral models. Under the passive bounded rationality model, individuals are thought to continue assessing all available attributes, however, do so with increasing levels of error as choice complexity increases (de Palma *et al.* 1994). The rationally-adaptive model assumes that individuals recognize that their limited cognition may have positive opportunity costs and react accordingly. As DeShazo and Fermo (2004) state:

> *"Individuals will therefore allocate their attention across alternative-attribute information within a choice set in a rationally-adaptive manner by seeking to minimize the cost and maximize the benefit of information evaluation" (page 3).*

---

[9] We are of the view that non-compensatory behaviour is confounded with attribute processing and that when one conditions the choice outcome on the heterogeneous set of AP rules, that compensatory behaviour is a good approximation. The real risk with non-compensatory choice models is that they are placing the 'not relevant' attributes at the lowest level in the EBA hierarchy without assessing whether they should be there at all.

[10] Significant research effort has been expended on how to optimise the outputs derived from respondents completing choice tasks derived from these single design plans, generated using statistical design theory (e.g., Bunch *et al.* 1994; Huber and Zwerina 1996, Kanninen 2002; Kuhfeld *et al.* 1994; Lazari and Anderson 1994; Sandor and Wedel 2001), whilst minimizing the amount of cognitive effort required of respondents (e.g., Louviere and Timmermans 1990; Oppewal *et al.* 1994; Wang *et al*. 2001, Bliemer and Rose 2005).

It is important to recognise that simplistic designs may also be ‘complex’ in a perceptual sense. Individuals may expect more information than was given to them, thinking such information would be relevant in a real market setting[11]. The development of a SC experiment, supplemented with questions on how an individual processed the information, enables the researcher to explore sources of systematic influences on choice. Examples of such questions are shown in the two screens below (Figures 2 and 3).

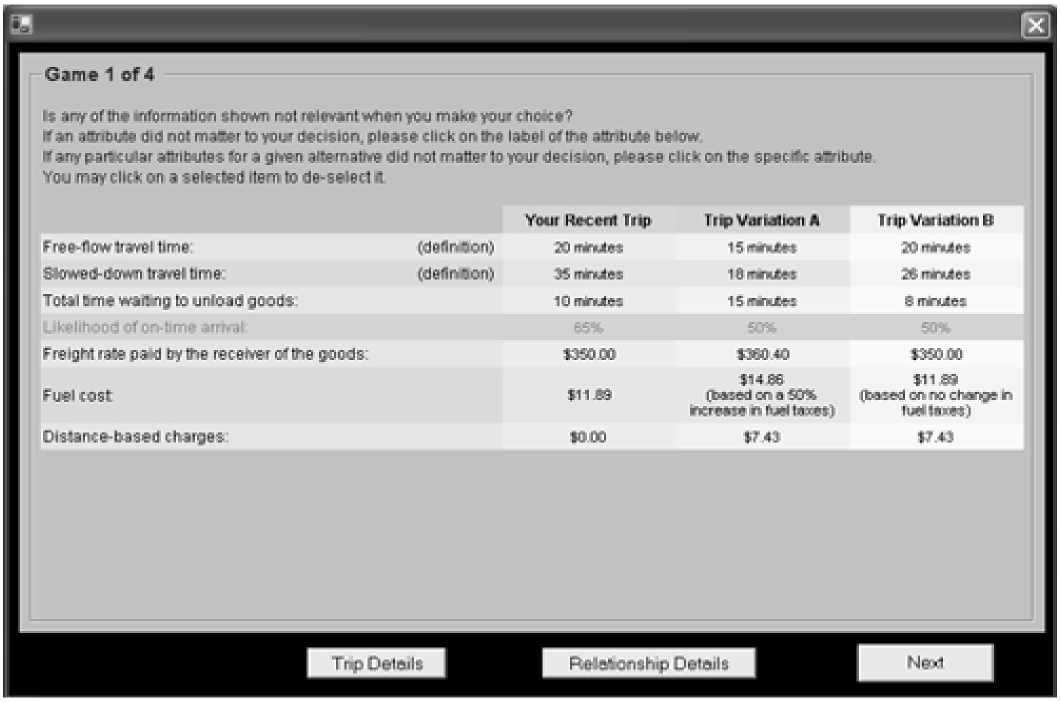


Figure 2: Attribute and alternative specific processing rules

[11] There is widespread evidence in the psychology literature on the behavioral variability, unpredictability and inconsistency regularly demonstrated in decision making and choices (e.g., Gonzales-Vallejo 2002; Slovic 1995), reflecting an assumption that goes back at least to Thurstone’s law of comparative judgment (1927). One of the particularly important advantages of using a stochastic representation of decision strategies, as promoted herein, is that it enables a more behaviorally realistic analysis of variation in decision strategies.

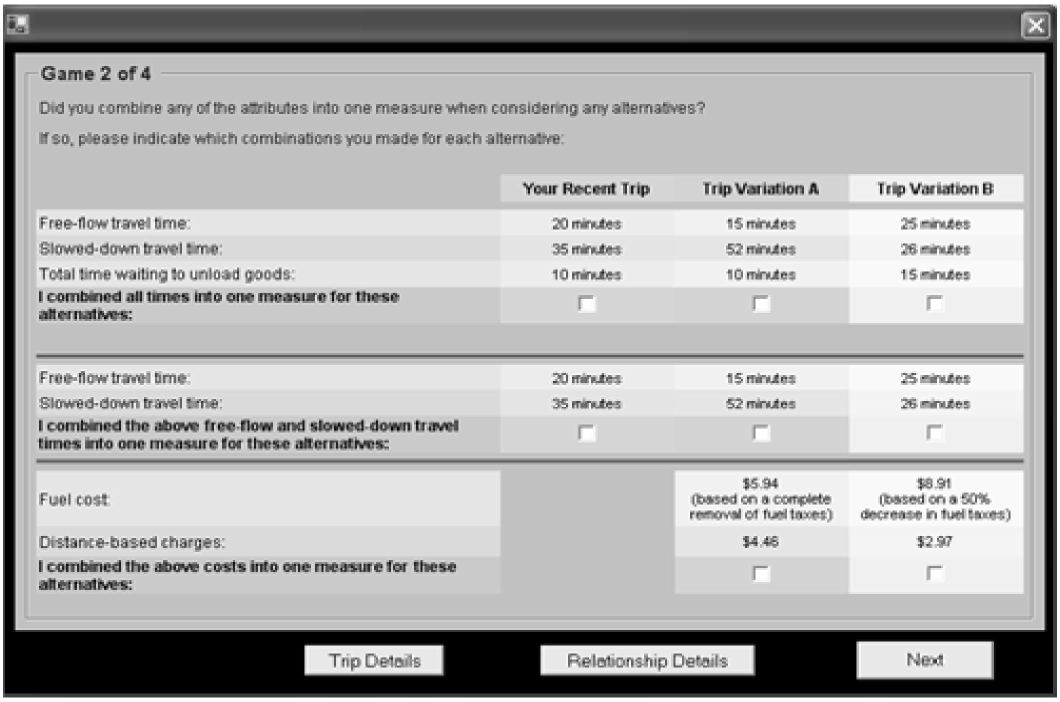


Figure 3: Inter-related attribute processing rules

There is a substantial extant literature in the psychology domain about how various factors affect the amount of information processed in decision tasks. Recent evidence demonstrates the importance of such factors as time pressure (e.g., Diederich 2003), cognitive load (e.g., Drolet and Luce 2004), and task complexity (Swait and Adamowicz 2001). There is also a great deal of variability in decision strategies employed in different contexts, and this variability adds to the difficulties of understanding the behavioral mechanisms. A recent attempt to define a typology of decision strategies (e.g., Payne, *et al*. 1992) has been particularly useful.

Payne *et al*. (1992) characterized decision strategies along three dimensions: basis of processing, amount of processing, and consistency of processing. Decision strategies are said to differ in terms of whether or not many attributes within an alternative are considered before another alternative is considered (*alternative-based* processing) or whether values across alternatives on a single attribute are processed before another attribute is processed (*attribute-based* processing). Strategies are also said to differ in terms of the amount of information processed (i.e., whether any information is ignored or not processed before a decision may be made). Finally, decision strategies can be grouped in terms of whether the same amount of information for each alternative is examined (*consistent* processing) or whether the amount of processing varies depending on the alternative (*selective* processing).

On the basis of this typology, Payne *et al*. (1992) identified six specific decision strategies, three of which are attribute-based and three alternative-based approaches. The attribute-based approaches included the elimination-by-aspects (EBA), lexicographic choice (LEX), and majority of confirming dimensions (MCD) strategies. The alternative-based approaches included the weighted additive (WADD), satisficing (SAT), and equal-weight (EQW) strategies. These strategies are further described in Table 1 below. The main argument posited by Payne *et al*. (1992) was that individuals construct strategies depending on the task demands and the information they are faced with.

Table 1: Typology of decision strategies

Strategy	Attribute or Alternative-based	Amount of Information	Consistency
EBA	Attribute-based	Depends on values of alternatives and cut offs	Selective
LEX	Attribute-based	Depends on values of alternatives and cut offs	Selective
MCD	Attribute-based	Ignores probability or weight information	Consistent
WADD	Alternative-based	All information processed	Consistent
SAT	Alternative-based	Depends on values of alternatives and cut offs	Selective
EQW	Alternative-based	Ignores probability or weight information	Consistent

The status quo in SC modeling is the WADD strategy, since it assumes that all information is processed. Elimination by aspects (See Starmer 2000) involves a determination of the most important attribute (usually defined as the attribute with the highest weight/probability) and the cut-off value for that attribute (i.e., a threshold). An alternative is eliminated if the value of its most important attribute falls below this cut-off value. This process of elimination continues for the second most important attribute, and so on, until a final alternative remains. Thus, the EBA strategy is best characterized as a *'threshold'* attribute processing strategy. The LEX strategy, in its strictest sense, involves a direct comparison between alternatives on the *most important* attribute. In the event of a tie, the second most important attribute is used as a comparison, and so on until an alternative is chosen. The LEX strategy is thus best characterized as a *'relative comparison'* strategy. Thus, we can clearly differentiate two classes of attribute processing strategies: threshold and relative comparison.

A major deficiency of these strategies is that although they assume selectivity in attribute processing across different decision task contexts, *they assume consistency in attribute strategy within the same decision context*. In other words, once a strategy is selected for a given task (or choice), it does not change within the task. This issue is further complicated by psychological theory which identifies two main stages in the decision process. Differentiation and Consolidation Theory, developed by Svenson (1992), assumes that decision-making is a goal-oriented task which incorporates the pre-decision process of differentiation and the post-decision process of consolidation. This theory is crucial in encouraging a disaggregation of the entire decision process.

The two issues discussed above, namely the *adaptive* nature of strategies and the *disaggregation* of the decision process, are issues that can only be assessed realistically within a paradigm that relaxes the deterministic assumption of most models of decision-making. This is consistent with the payoff-probability structure discussed above. A preferred approach would involve a stochastic specification of attribute processing that is capable of accommodating the widespread consensus in the literature that decision-making is an active process which may require different decision making strategies in different contexts and at different stages of the process (e.g., Stewart *et al.* 2003). As the relevance of attributes in a decision task changes so too must our approach to modeling the strategies individuals employ when adapting to such changes. Specifically we need a flexible framework within which we can accommodate the influence of one or more of the processing strategies on choice making across the sampled population.

2.3 How do analysts account for heterogeneous attribute processing?

How is the attribute processing strategy (APS) of each individual best represented within the SC modeling framework? The editing stage of prospect theory (see Starmer 2001, Kahnemann and Tversky 1979) is a useful theoretical setting; in this stage, agents use heuristics to make a decision setting optimally tractable. The APS can be partitioned into: (i) processes associated with decision making in real markets, and (ii) processes invoked to accommodate the information load introduced by the SC survey instrument.

Hensher (2004) has shown that the two processes are not strictly independent. The processing of an SC experiment has some similarity to how individuals process information in real markets[12]. The APS may be hypothesized to be influenced by relevant information sources resident in the agent's memory bank, either processing instructions or knowledge sources. Specific processing instructions can include: (i) reference dependency[13], (ii) event and attribute splitting, (iii) attribute re-packaging, (iv) the degree of information preservation, and (v) the role of deliberation attributes. Knowledge sources can include the macro-conditioners.

We can view the treatment of process via one or more rules, as a deterministic or stochastic specification. In Hensher *et al.* (2005), for example, we treated the exogenous information of attribute inclusion/exclusion deterministically. We assumed that the analyst knows for certain which attributes are used by which respondents. It is probably more realistic, however, for the exogenous information to point to the correct likelihood specification, so that the likelihood for a respondent is a probabilistic mixture of likelihoods (Hensher *et al.* in press). In contrast to a deterministic specification, which assumes knowledge of the respondent-level likelihood of attribute processing with certainty, a stochastic specification relaxes this assumption. One way of defining a stochastic model is to assume that the exogenous covariate is probabilistically related to the structural heterogeneity specification, through an expected maximum utility index derived from a choice of attribute processing strategy model, conditioning the preference heterogeneity distribution for each random parameter associated with the attributes of the SC model.

To illustrate this point, using a sample of car non-commuters in Sydney we estimated a mixed logit model in which all attributes are assumed to be attended to, and models which assume that certain attribute(s) are not attended to, based on supplementary information provided by respondents (see Table 2). The supplementary information is

[12] The main difference is that the SC experiment provides the information to be processed, in contrast to real markets where more effort is required to search for relevant information. We recognise, however, that the amount of information in the SC experiment may be more than what an individual would normally use in making a choice. Yet that is precisely why we have to establish the APS of each individual to ensure that the offered information is represented appropriately in model estimation. For example, if an attribute is ignored, we need to recognise this and not assume it is processed as if it is not ignored.

[13] This is defined empirically by the relative distance between the attribute levels in the SC alternative and levels that an individual is familiar with (i.e., a case-based-decision-theoretical memory set that actually has been experienced as defined herein by the base alternative – a recent or a most-experienced alternative). Reference dependency is a member of the broader class of the similarity condition of CBDT in which it is suggested that individuals choose acts based on their performance in similar problems in the past. The review and assessment of a choice task is defined as a *problem* in CBDT.

accounted for in a deterministic and a stochastic way; the latter in recognition of the analyst's lack of full information on why a specific attribute processing (AP) strategy was adopted by each sampled individual. We compare the value of travel time savings distribution under alternative attribute processing regimes (Table 3). As expected, there are significant variations in the mean and standard deviation willingness to pay (WTP) across the three AP strategies.

Table 2: Utility expressions for attribute attention profiles, estimated as multinomial logit

	Attribute Processing Profile
V_1	All attributes attended to
	Attributes not attended to:
V_2	Running cost
V_3	Running and toll cost
V_4	Toll Cost
V_5	Slowed down time
V_6	Free flow and slowed down time
V_7	Free flow time
V_8	Slowed down time and running cost
V_9	Free flow and slowed down time and to cost

$V_1=2.0909+0.02872\times \text{age}-0.01088\times \text{income}-0.03606\times \text{ff}+0.11071\times \text{sdt}+0.1969\times \text{cost}+0.06767\times \text{toll}$
$V_2=1.7487+0.019159\times \text{age}-0.011466\times \text{income}-0.03545\times \text{f}+0.10151\times \text{sdt}+0.17557\times \text{cost}+0.06932\times \text{toll}$
$V_3=-1.49000+0.01978\times \text{age}-.001379\times \text{income}\mathbf{-.00194\times ff}+0.13364\times \text{sdt}\mathbf{+0.07899\times cost+0.01865\times toll}$
$V_4=-3.055\mathbf{+0.01147\times age}+0.01349\times \text{income}-0.020047\times \text{ff}+0.1175\times \text{sdt}+0.20619\times \text{cost}+0.07678\times \text{toll}$
$V_5=0.82309+0.03845\times \text{age}-0.01994\times \text{income}\mathbf{-0.01032\times ff}-0.05525\times \text{sdt}+0.33109\times \text{cost}\mathbf{+0.00305\times toll}$
$V_6=1.68608+0.01397\times \text{age}-0.02204\times \text{income}-0.061966\times \text{ff}+0.126399\times \text{sdt}+0.2674\times \text{cost}+0.0999\times \text{toll}$
$V_7=1.5842-0.02523\times \text{age}\mathbf{-0.003078\times income}-0.017136\times \text{ff}+0.07665\times \text{sdt}+0.14232\times \text{cost}\mathbf{-.016056\times toll}$
$V_8=-4.10832+0.07469\times \text{age}-.0112178\times \text{income}-.03349\times \text{ff}+.12575\times \text{sdt}+.23752\times \text{cost}\mathbf{-.00806\times toll}$
$V_9=0$
Pseudo-R^2 = 0.179, bolded= statistically non-significant at 95 percent confidence level

Table 3: Values of travel time savings ($ per person hour car non-commuter driver)

time = random parameter, cost = fixed parameter

Attribute	All attributes assume to be attended to		Deterministic attribute exclusion		Stochastic attribute exclusion	
	mean	Std dev	mean	Std dev	mean	Std dev
Free flow time	7.60	0.47	7.81	0.46	7.95	3.59
Slowed down time	9.33	0.57	10.65	0.67	9.91	1.22
Ratio slowed to free flow time	1.23	1.21	1.36	1.46	1.25	0.69
Confidence level (95%):						
Free flow time	0.02		0.02		0.12	
Slowed down time	0.02		0.02		0.04	
Sample Size	3568		3071/2944*		3568	

* 3,071 r×elates to free flow and 2,944 relates to slowed down time.

Defining the choice set of AP strategies is also important and is a little-studied issue. Hensher (2004) investigated one AP strategy, where the alternatives were defined as the number of preserved attributes (0, 1, 2, …). This is appealing in the sense that an individual, when evaluating alternatives in a choice set, as defined by a set of attributes, has in front of them information from the attributes (number, levels and range) that varies across the alternatives. The individual then processes this information by invoking a series of rules that appear to be linked to the processing instructions given above. Given the central role of a SC experiment in the parameterisation of the utility expressions that describe preference formation and equilibrium choices up to a probability of choice, the APS alternatives might reasonably be defined by the dimensionality of each choice task. Parameterisation of the APS alternatives will reveal the sources of information brought to bear on the way that individuals establish their preferences for specific alternatives[14].

Since the choice made by an individual is conditioned on the APS, and given the two-stage decision process promoted in prospect theory, it is desirable to re-specify the choice model as a two-stage processing function, wherein each individual's choice of alternative is best represented by a joint choice model involving the individual's choice conditional of the APS and the (marginal) choice of APS (Figure 4). We then have to decide which set of influences reside in the APS utility expression and in the choice utility expression. We anticipate that it is the processing rules that reside in the APS expressions (e.g., equation 1) and the attributes of alternatives that reside in the choice utility expressions. The contextual and person-specific interactions may reside in both sets of utility expressions. The APS utility expression might be, in a linear form (although non-linearity should be tested):

$$U_{aps\_i} = \alpha + \beta_1 AddAtts_i + \beta_2 \#IgnAtts_i + \beta_3 RefDepX_{1i} + \beta_4 RefDepX_{2i} + \beta_5 IV_i \quad (1)$$

where IV_i is the expected maximum utility associated with the choice process at the lower level of the tree structure proposed in Figure 4, similar to the theoretical link established within a nested logit model. This model recognizes that the APS is influenced by the actual information setting within which the preferred contract outcome is selected by an agent.

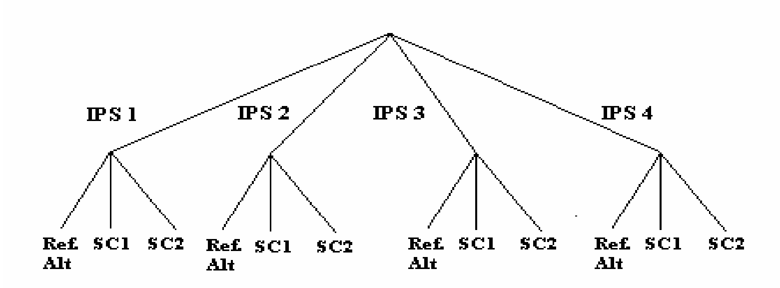


Fig 4: Individual-specific decision structure for SC assessment

[14] Importantly, in order to establish the full dimensionality of an agent's APS, we must show them the full attribute design and establish how they choose to process it. This is essential for each choice set if we are to assess the influence of reference dependency as defined by the levels of attributes in each SC alternative relative to the reference alternative (i.e., experienced or memory-based) alternative.

The approach described above implies a specific experimental design strategy. All individuals are given a single design specification in terms of the constituent attribute dimensions (number of attributes, number of levels of each attribute, attribute range) plus a fixed number of alternatives. For each choice task, a choice is made and then supplementary questions establish how the choice task is processed in terms of the invoking of one or more of the processing instructions listed above.

Alternatively, we might establish the APS more directly through the first stage of a two-stage choice experiment. In stage 1 we might offer a number of pre-designed choice experiments with varying numbers, levels and range of attributes across two alternatives, plus a reference alternative (from the agent's memory bank). These attributes can be structured in each design (in accordance with D-optimality conditions of experimental design – see below) under rules of preservation, attribute re-packaging and relativity to the reference alternative[15]. Individuals would be asked to evaluate each design and to indicate their preferred design in terms of the information that matters to them (i.e., relevancy). We could then identify, across all designs, what information is irrelevant for behavioral processing and what is ignored to avoid cognitive burden. We can also establish the extent to which specific alternatives are seen as similar to prior accumulated experience resident in the memory bank of the individual, which are recalled as an aid in AP (since this links nicely to the notion of similarity-weighted utility in choice-based decision theory)[16].

Hensher *et al.* (2005a) and Hensher and Pucket (2005) have implemented the APS choice method in the context of urban freight distribution in a supply chain, where transporters and shippers were interviewed. Identification of the role of different attribute processing strategies in a model of choices amongst attribute packages (as shown in Figure 1 above) is elicited through Figures 2 and 3. Tables 4 and 5 summarise the degree to which attributes in the model were assigned an adjusted value for marginal (dis)utility through the adoption of specific APSs.

[15] The range of possible APS's would be established in prior in-depth interviews with stakeholders. The advantage of this two-stage approach is that *each design* (conditioned on the APS) will be D-efficient.

[16] Establishing how similarity from memory is built into the estimation of the choice model is challenging. As a global condition throughout the utility expression, it can be treated as an exogenous adjustment through a discrete-continuous choice specification. For example, we might estimate a similarity model where the dependent variable is some measure of 'similarity', and then use the predicted similarity indicator as a multiplicand of the utility estimate attached to each alternative in the discrete choice model prior to deriving the choice probabilities.

Table 4: Occurrence of attribute exclusion

Attribute	Number of Times Ignored (Frequency)
Free-flow time (transporter)	43 (10%)
Free-flow time (shipper)	224 (27%)
Slowed-down time (transporter)	53 (12%)
Slowed-down time (shipper)	268 (33%)
Waiting time (shipper)	324 (40%)
Probability of on-time arrival (transporter)	49 (11%)
Probability of on-time arrival (shipper)	61 (7%)
Freight rate (transporter)	40 (9%)
Freight rate (shipper)	82 (10%)
Fuel cost (transporter)	29 (7%)
Fuel cost (shipper)	242 (30%)
Variable charges (transporter)	25 (6%)
Variable charges (shipper)	246 (30%)

Table 5: Occurrence of attribute aggregation

Attribute	Number of Times Aggregated (Frequency)
Time measures (transporter)	292 (68%)
Time measures (shipper)	271 (33%)
Cost measures (transporter)	326 (75%)
Cost measures (shipper)	341 (42%)

To establish the influence that the distribution of attribute processing strategies has on a key behavioural outputs such as the value of travel time savings (VTTS), we estimated mixed logit models (reported in Hensher *et al.* 2005a) for the non-APS (Figure 1 only) and APS (Figures 1-3) data. The estimation sample includes 108 transporters and 102 shippers, yielding 1,248 observations (432 choice sets faced by transporters and 816 choice sets faced by shippers). Table 6 summarises the variation in VTTS measures across attribute exclusion and aggregation strategies for transporters, to contrast with the findings under a non-APS model for transporters (Table 7).

Table 6: VTTS Measures (AU$ per hour) for APS models

Note: exclusion has been accounted for prior to the aggregation condition

	Agg. Time (Agg. Cost)	**Agg. Time (Dis. Costs)**	**FF (Agg. Cost)**	**SDT (Agg. Cost)**	**FF (Dis. Costs)**	**SDT (Dis. Costs)**
Proportion of Sample	64.4%	5.6%	18.4%	18.4%	11.7%	11.7%
Mean	$19.36	$37.66	$64.38	$84.53	$134.82	$178.41
Std. Deviation	$18.69	$30.83	-	--	$82.01	$108.42
Minimum	-$55.63	-$38.44	--	--	$75.92	$99.69
Maximum	$87.57	$102.16	--	--	$274.53	$360.48
Proportion of Neg. Values	4.8%	5.6%	0%	0%	0%	0%

Table 7: VTTS measures (AU$ per hour) for non-APS model

	Free-Flow Time	**Slowed-Down Time**
Mean	$42.48	$83.77
Standard Deviation	$22.95	$8.88
Minimum	-$22.64	$55.67
Maximum	$99.39	$162.42
Proportion of Negative Values	1.9%	0%

When taking attribute exclusion and aggregation into account, strong variation in VTTS estimates is found across AP strategies. An assumption of passive bounded rationality assigns a uniform relationship between free-flow and slowed-down time savings across the sample, whilst the APS model allocates significantly different ratios in VTTS measures for free-flow time and slowed-down time across exclusion and aggregation strategies. The most straightforward case involves the aggregation of time measures, which, by definition, results in no difference in the valuation of a unit of free-flow time versus a unit of slowed-down time. This lack of variation in VTTS for a given decision maker across time components is countered by differential willingness to pay between the valuation of free-flow and slowed-down time for those who distinguished between the two measures. For both those who aggregated cost measures and those who kept them separate, the ratio of VTTS for slowed-down time to VTTS for free-flow time is approximately 1.32.

This ratio is tempered relative to that found in the non-APS model (Table 7), in which the VTTS for slowed-down time is almost twice the VTTS for free-flow time. There are two highly significant implications of this discrepancy: (1) the inclusion of APS information into the model results in a lower inferred premium placed by transporters on the mitigation of slowed-down time; and (2) heterogeneity in processing strategies for costs does not have an impact on this relationship. In other words, acknowledging the aggregation strategies of transporters with respect to time has a significant impact on the resulting behavioral implications with respect to time savings; furthermore, acknowledging the aggregation strategies of transporters with respect to cost does not obscure this relationship at all.

The utilization of APS information in model estimation identifies sub-groups, each of which holds a distinct distribution of VTTS. The choice not to differentiate between free-flow and slowed-down time results in the presence of no unique disutility of slowed-down time for those who opted to aggregate time measures. However, the link between this aggregation strategy and VTTS goes beyond the direct relationship between free-flow and slowed-down time: transporters who aggregated transit time measures appear to value travel time savings much lower than those who treated free-flow and slowed-down time separately, regardless of cost aggregation strategy. The mean VTTS estimates for those who aggregated time measures are considerably lower than the remainder of transporters, at only \$19.36 and \$37.66 per hour when aggregating costs and keeping costs separate, respectively.

These values, which are less than half of the mean estimates in the non-APS model, are in stark contrast to the mean VTTS estimates for those who did not aggregate time measures. The sub-group who aggregated costs but not times has a mean VTTS that is close to the mean estimates from the non-APS model, after considering the difference in ratios of free-flow and slowed-down VTTS across the models. That is, the mean VTTS for slowed-down time is very similar in the two models (\$83.77 per hour in the non-APS model versus \$84.53 in the APS model), whilst the mean VTTS for free-flow time conforms to the general ratio between free-flow and slowed-down VTTS in each model.

Most strikingly, respondents who attended to all time and cost measures individually demonstrate a high VTTS for free-flow and slowed-down time, with mean values above the corresponding maximum values in the non-APS model. Whereas the non-APS model identified the presence of some VTTS estimates well above the mean, these values could be interpreted as artefacts of the distributional assumptions on the random parameters. However, upon including APS information into the modeling process, one finds that a small proportion of transporters (the 11.7 percent who kept times and costs disaggregate) do have much higher values of travel time savings, relative to the remainder of the sample.

This evidence clearly shows that there are systematic (i.e., attribute processing) forces driving the variation in behavioral measures. Indeed, it may be the case that attribute processing strategies serve as proxies for factors that may be difficult to otherwise capture, such as the profitability of respondents, or their flexibility in utilising the efficiency gains offered through variable charges. That is, relatively low or high VTTS measures may be indicative of both the ability of the respondent’s organization to take advantage of efficiency gains (e.g., utilizing the truck in an additional task that would not be otherwise possible), and the magnitude of net benefits afforded through these efficiency gains (i.e., the net profitability of any potential subsequent freight task that is made available through gains in travel quality).

3. Joint decision making and identifying cooperation amongst agents

In any negotiation between two or more agents, they each bring to a negotiation a set of preferences. These preferences are assumed to be consistent with an agent-specific utility maximization rule. Through negotiation however, each participating agent may have to make compromises to arrive at set of agreements that will enable the group as a whole to fulfill its joint objective.

One model system that has appeal as a way of establishing the preferences of each agent, and the role that each agent's individual preferences play in establishing the group preference function is summarised below.

Stage 1: Each agent participates in a SC experiment with common choice sets but with permissible different attribute processing strategies. The behavioural process assumes that each agent acts *as if* they are a utility maximiser. The agent-specific models define utility expressions of the form: U(alt *i*, agent *q*) i=1,…,J; q=1,…,Q where *alt* defines an alternative outcome (or choice). For example, with two agents and three alternative outcomes we have $U(a_1q_1)$, $U(a_2q_1)$, $U(a_3q_1)$ for agent 1 and $U(a_1q_2)$, $U(a_2q_2)$, $U(a_3q_2)$ for agent 2. A labelled or unlabelled SC design can be established to parameterise this independent-utility-maximising choice model, conditional on the APS of the agent.

The relative attribute preservation of the APS of the agent is identified by prompting agents to indicate the attributes that were ignored or given little attention for each outcome alternative. Questions about the APS enacted by agents within each choice set can be used to test a range of APS choice models of varying complexity (see Hensher 2004, DeShazo and Fermo 2004). This information is used to condition the utility expressions as per section 2.

The base utility expressions (i.e., without any interaction effects or direct covariate effects) are of the general form:

$$U_{qj} = \alpha_j + \boldsymbol{\beta}_{qjk} \times \mathbf{x}_{qjk} + \varepsilon_{j},, \tag{2}$$

where $\mathbf{x}_{qjk}$ is a vector of design attributes associated with agent *q* and alternative *j*, $\boldsymbol{\beta}_{qjk}$ is the corresponding vector of random marginal utility parameters, α_j is an alternative-specific constant and ε_j represents the unobserved effects. The effect of the APS used by an agent for a given choice set is implemented by setting $\boldsymbol{\beta}_{\mathbf{qjk}} = 0$ if *k* is ignored for a particular alternative *j* for agent *q*. The mean and the standard deviation of the random preference parameters $\boldsymbol{\beta}_{\mathbf{qjk}}$ across the sample of agents can both be decomposed, and hence explained, by deliberation attributes such as the number of years involved in the specific decision setting, socioeconomics characteristics; prior experience in such negotiations, and general APS-related information such as the number of attributes *ignored*[17]. Regardless of which approach is adopted, such contextual influences can also be interacted with design attributes in model estimation. This modelling structure lends itself to the heterogenous mixed logit (HML) model, which is our econometric model of choice for this methodology.

Stage 2: All parameters estimated from stage 1 are fixed and imported[18] into a joint agent model. For example, with two agents and three alternatives, there are nine joint alternative propositions – $U(a_1a_1)$, $U(a_1a_2)$, $U(a_1a_3)$,…., $U(a_3a_1)$, $U(a_3a_2)$, $U(a_3a_3)$, referred to as propositions p=1,…,P. Three of the joint propositions imply non-negotiated cooperation (i.e. U(a1a1), U(a2a2) , U(a3a3)). The stage 2 choice is between combinations of agent-specific propositions with one proposition the chosen pair. A model is then specified of the following form (for two agents, q, _q):

[17] Treated either deterministically or stochastically.

[18] The alternative-specific constants may not be imported.

$$U(a_1a_1) = ASC_{a1a1} + \lambda_{qp} \times (\beta_{1q}x_{1q} + \beta_{2q}x_{2q} + ...) + (1-\lambda_{qp}) \times (\beta_{1\_q}x_{1\_q} + \beta_{2\_q}x_{2\_q} + ...)$$

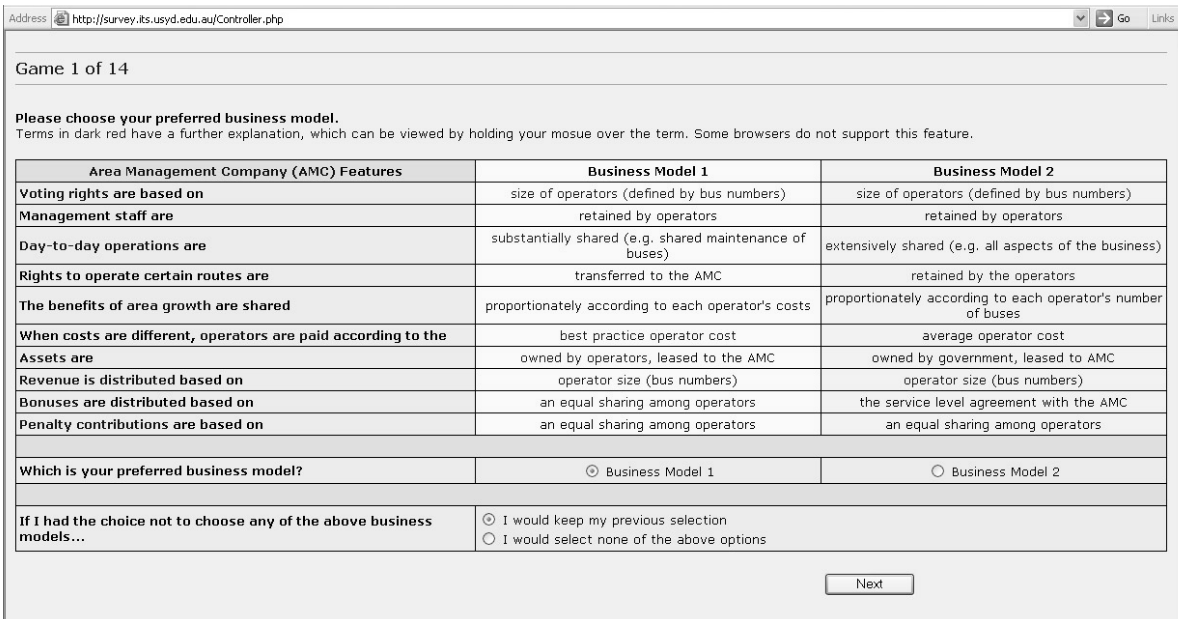
$$U(a_1a_3) = ASC_{a1a3} + \lambda_{qp} \times (\beta_{1q}x_{1q} + \beta_{2q}x_{2q} + ...) + (1-\lambda_{qp}) \times (\beta_{1\_q}x_{1\_q} + \beta_{2\_q}x_{2\_q} + ...) \quad (3)$$

$$U(a_3a_3) = ASC_{a3a3} + \lambda_{qp} \times (\beta_{1q}x_{1q} + \beta_{2q}x_{2q} + ...) + (1-\lambda_{qp}) \times (\beta_{1\_q}x_{1\_q} + \beta_{2\_q}x_{2\_q} + ...)$$

The power measures for agents *q* and *_q* sum to one, making comparisons of agent types straightforward. If the two power measures are equal for *a given attribute mix defining a proposition* (i.e., λ_{qp} = (1 - λ_{qp}) =.5), then group choice equilibrium is not governed by a dominant agent with respect to proposition *p*. In other words, regardless of the power structure governing other attributes, agent types *q* and *_q* tend to reach perceptively fair compromises when bridging the gap in their preferences for each proposition. If the power measures are significantly different across agent types (e.g., λ_{qp} > (1 - λ_{qp}) for two agents), then λ_{qp} gives a direct measure of the dominance of one agent type over the other with respect to attribute mix in proposition *p*; as λ_{qp} increases, so does the relative power held by agent type *q* over *_q*. For example, the power measures may reveal that one agent type tends to get its way with regard to monetary concerns, whereas the other agent type tends to get its way with regard to concerns for levels of service. These relationships can be examined further at the sub-type level (by decomposition of the random parameter specification of λ), in order to reveal deviations from the inferred behaviour at the sample level that may be present for a particular type of relationship.

This model is straightforward to estimate, holding all β's fixed, with each λ_{qp} and the alternative-specific constants (ASC's) as free parameters. λ_{qp} as a power indicator can be a random parameter and a function of other criteria, especially the deliberation attributes, and can be specific to each attribute within and/or between propositions, or constrained as the analyst sees fit.

To illustrate the empirical appeal of this method, we draw on a study of negotiation between bus operators in Sydney who, from January 2005, w required to form alliances with one of more other incumbent operators in a spatially realigned contract regime, where 37 contract areas are reduces to 15 (see Hensher and Knowles 2005 for more details). The data used to estimates the models in Stages 1 and 2 is obtained from an internet-based SC experiment (see Figure 5 for an example of a SC screen).



Address http://survey.its.usyd.edu.au/Controller.php Go Links

Game 1 of 14

Please choose your preferred business model.
Terms in dark red have a further explanation, which can be viewed by holding your mosue over the term. Some browsers do not support this feature.

Area Management Company (AMC) Features	Business Model 1	Business Model 2
Voting rights are based on	size of operators (defined by bus numbers)	size of operators (defined by bus numbers)
Management staff are	retained by operators	retained by operators
Day-to-day operations are	substantially shared (e.g. shared maintenance of buses)	extensively shared (e.g. all aspects of the business)
Rights to operate certain routes are	transferred to the AMC	retained by the operators
The benefits of area growth are shared	proportionately according to each operator's costs	proportionately according to each operator's number of buses
When costs are different, operators are paid according to the	best practice operator cost	average operator cost
Assets are	owned by operators, leased to the AMC	owned by government, leased to AMC
Revenue is distributed based on	operator size (bus numbers)	operator size (bus numbers)
Bonuses are distributed based on	an equal sharing among operators	the service level agreement with the AMC
Penalty contributions are based on	an equal sharing among operators	an equal sharing among operators
Which is your preferred business model?	Business Model 1	Business Model 2
If I had the choice not to choose any of the above business models...	I would keep my previous selection I would select none of the above options	

Next

Fig 5: An example of a stated preference screen

A total of 19 operators completed the online pilot survey, yielding 266 observations for model estimation (i.e., 14 choice sets by 19 bus operators). There were only two statistically significant effects – payments (where costs differ) based on best practice costs, and assets owned and operated by each operator within the AMC. For best practice costing, the distinction between metropolitan and non-metropolitan operator had an influence. The two significant effects are defined by random parameters, suggesting that preference heterogeneity is relevant. A constrained triangular distribution was selected as the preferred analytical distribution (from a number of other distributions assessed).

Operators have a strong positive preference for owning and operating their own assets within an AMC and for being paid on the basis of best practice costs where operator costs differ. Furthermore the preference heterogeneity is most marked, as shown by the distributions in Figures 6 and 7. For best practice costing, the preference level has been conditioned on whether the operator is metropolitan or not. The positive parameter estimate for the decomposition of the mean (i.e., 1.0074), suggests, all other influences remaining constant, that the marginal utility associated with best practice costing is higher for metro operators than for non-metro operators. The range of marginal utilities varies from 0.623 to 0.844 for asset ownership, and from 0.355 to 1.383 for best practice costs. Hence the preference heterogeneity, while significant for both attributes, is much greater for best practice costing. The mean for each attribute is respectively 0.766 and 0.704, highlighting the potential for misleading inferences when reliance is on the mean of a distribution with a wide range of marginal utilities. The respective standard deviations are 0.059 and 0.476.

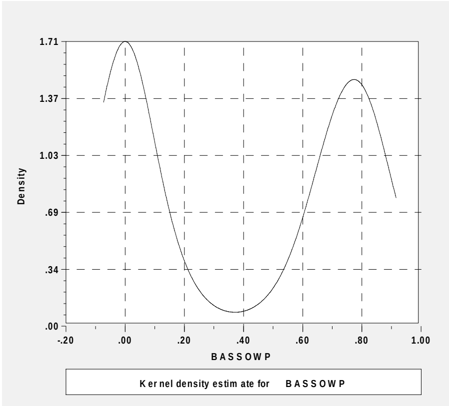


Figure 6: Asset ownership preference profile

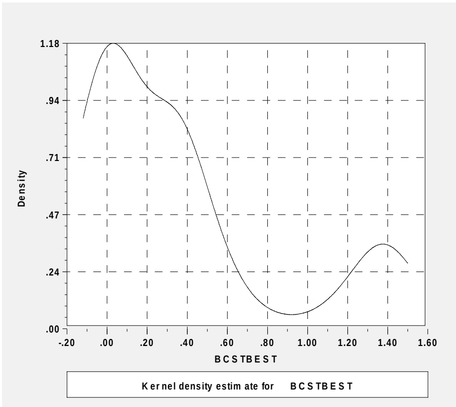


Fig 7: Best practice costing preference profile

One objective is to use the findings for each operator to establish the extent to which they would cooperate *if they were required* to work together under a single AMC. Essentially the cooperative spectrum is based on matching of aggregate utility from specific business propositions. Figure 8 profiles the sample distribution of total utility associated with each of the three business propositions relative to a base of having assets not owned and operated by each operator and zero for not using best practice costing. The indexation in Tot U_{ij} refers to the preference for ownership and operation of assets by operator (i=1) and best practice costing (j=1).

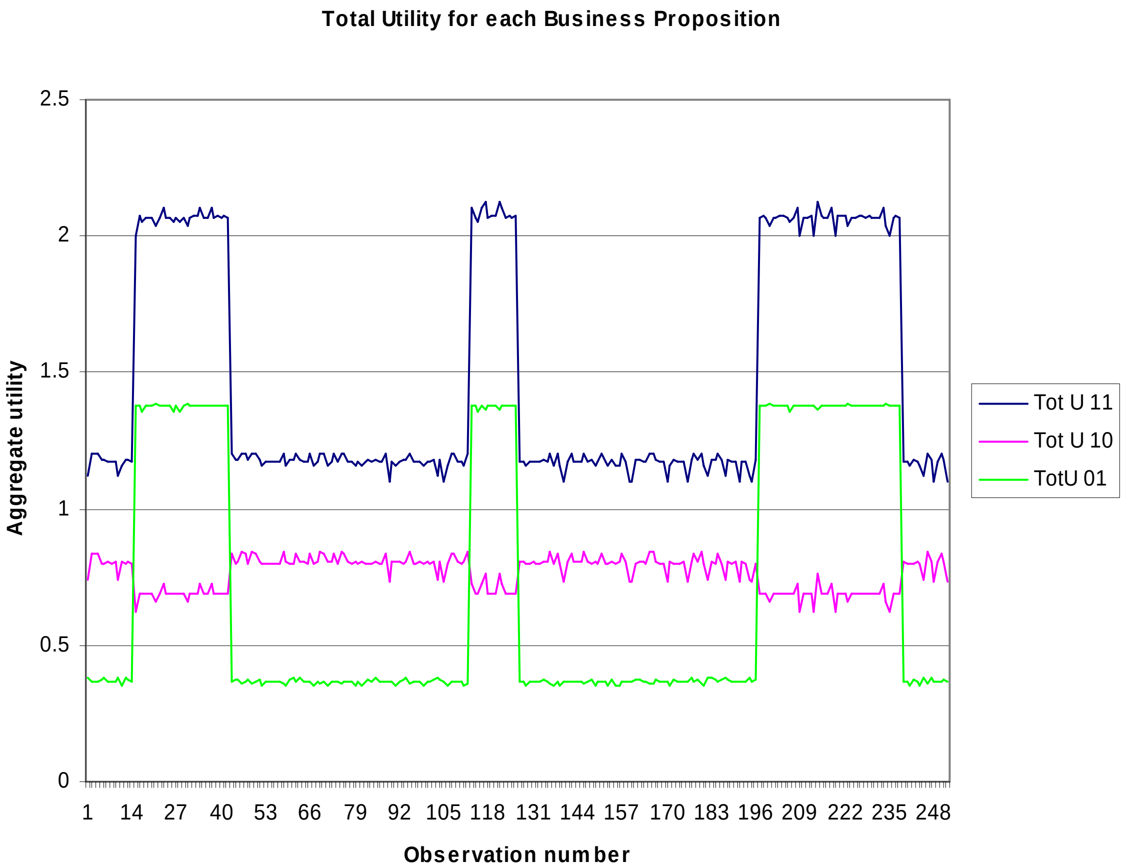


Fig 8: Profile of aggregate utility for each AMC proposition

The utility profiles in Figure 8 would be fed into equation (3) and the power weights estimated on all members of a specific AMC setting. We have done this for the sample and run two models – a simple multinomial logit (MNL) with a fixed parameter for the power or cooperation weight, and a mixed logit (ML) model in which the weight is a random parameter. We found that the cooperation weight (or lambda), treated as generic across all four AMC propositions, has a mean estimate of 0.529 (asymptotic *t*-value =3.705) from the MNL model and a mean of 0.516 (asymptotic *t*-value = 3.628) from the ML model. We used a constrained triangular distribution in which the standard deviation of the parameter is the same as the mean. The cooperation weight distribution is shown in Figure 9. Thus on average, the value of 0.5 suggests that any pair of parties appear to bring equal influence to the table (given the randomized pairwise matching we undertook herein to illustrate the types of useful outputs).

Of greater interest however is the cooperative strength on each of the four preference profiles in terms of best-practice cost and ownership and operation of assets. We find the following MNL and ML results (Table 8). There is clearly greater cooperation when both parties prefer to adopt best practice costing and well as maintain ownership and operations of ones' own assets; with least cooperation (albeit not statistically significant) when both prefer best practice costing, but disagree on asset ownership and operations. The MNL and ML results are similar, which might be expected given the size of the sample and the limited attribute influences. This reveals barriers to cooperation and gives some focus to where the parties need to dialogue more in arriving at a cooperative partnership.

Table 8: Cooperation (power) weights under alternative attribute preference profiles

Assets owned/operated by operator	Best practice costing	MNL – fixed parameter	ML – random parameter
yes	Yes	0.587 (2.8)	0.587 (2.80)
yes	No	0.423 (1.6)	0.439 (1.58)
no	Yes	0.230 (0.86)	0.247 (0.89)
no	No	0.625 (3.15)	0.631 (3.07)

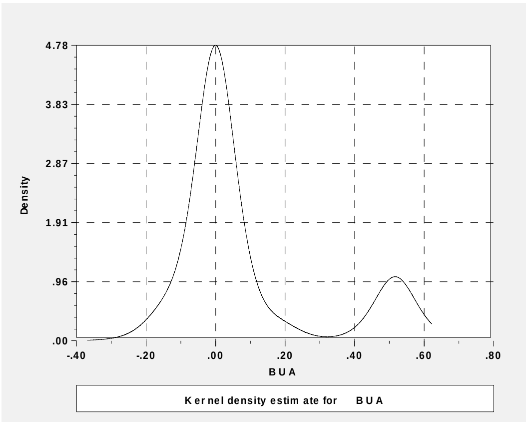


Fig 9: Cooperative (power) weight distribution

A *Stage 3* can be introduced which emphasizes a concession strategy (if there is suitable information, such as the response set in Figure 1 above) whereby one of the two agents has offered to concede to the preferences of the other party. This results in what we refer to as a concession model. With two agents, in the context of each agent making a separate assessment, we can establish the bounds of a direct negotiation outcome (i.e., a true IACE), where each bound reflects agent 1 concession and agent 2 concession. Details of this application are given in Hensher *et al.* (2005a) and Hensher and Puckett (2006).

4. Stated choice experimental design strategies

Conceptually speaking, an experimental design may be viewed as nothing more than a matrix of numbers that are used to assign values to the attributes of the alternatives present within the hypothetical choice situations of SC surveys (such as that shown in Figure 1). Typically, the allocation of the levels shown in these hypothetical choice situations are predetermined, systematically drawn from some underlying experimental design. For example, the attribute level values shown in Figure 1 are related to the attribute levels of a design, *x*, associated with each of the alternatives *j*, which may differ for each individual, *n*, as well as over each choice situation, *s*. By using experimental design theory, the assignment of these values occurs in some systematic (i.e., non-random) manner. What determines the systematic processes underlying the assignment of attribute level values to choice situations will be the basis of discussion in this section of the paper.

A cursory examination of the transportation literature suggests that the majority of SC studies applied to transportation contexts employ orthogonal fractional factorial designs[19] as opposed to using designs generated using *efficient* design techniques. The two approaches differ in that orthogonal fractional factorial designs are simply designs in which the attribute levels are orthogonal (uncorrelated) whereas efficient design techniques *typically* seek to generate designs which are not necessarily orthogonal but which minimize the asymptotic standard errors (and hence maximize the asymptotic *t*-ratios) of the parameter estimates to be obtained from a design. Independent of the type of design employed, experimental designs underlying SC studies require that respondents be shown one or more choice situations consisting of alternatives, each defined by a number of attributes which take discrete values called attribute levels.

4.1 Efficient designs

A statistically efficient design is a design that minimizes the elements of the asymptotic (co)variance matrix, Ω, with the aim of producing greater *reliability* in the parameter estimates given a fixed number of choice observations. In order to be able to compare the statistical efficiency of SC experimental designs, a number of alternative approaches have been proposed within the literature (see e.g., Bunch *et al.* 1994). The two most commonly used measures found within the literature are those of A-error and D-error.

$$\text{A-error} = \left(\text{trace}\ \Omega\right)^{k} = -\frac{1}{N}\left(\text{trace}\left(\frac{\partial LL(\beta)^2}{\partial\beta\partial\beta'}\right)\Big/ k\right) \tag{4}$$

$$\text{D-error} = \left(\det\Omega\right)^{1/k} = -\frac{1}{N}\left(\det\left(\frac{\partial LL(\beta)^2}{\partial\beta\partial\beta'}\right)\right)^{-1/k}. \tag{5}$$

Where k represents the number of parameters for the design, $LL(\beta)$ the log-likelihood function of the discrete choice model under consideration, N the sample size (we discuss the role sample size plays in generating efficient SC experiments below), and β the parameters to be estimated from the design. Given that we are generating designs and not estimating parameters for an already existing design, it is necessary to assume a set of priors for the parameter estimates. Given uncertainty as to the actual population parameters, it is typical to draw these priors from Bayesian distributions rather than assume fixed parameter values. Typically normal and uniform Bayesian distributions are used (uniform distributions are used if the direction and magnitude of the parameter estimates are unknown; e.g., Kessel *et al.* 2006). When Bayesian priors are assumed, the A and D-error measures are referred to as A_b and D_b error (where subscript b means Bayesian).

[19] Unfortunately, a large number of studies do not provide any information as to the type of design being used, nor the properties of the designs.

The $A_{(b)}$-error is computed by taking the trace of the asymptotic (co)variance matrix, whilst the $D_{(b)}$-error is calculated by taking the determinant, with both scaled to take into account the number of parameters to be estimated. The trace of a matrix is calculated as the sum of the diagonals of that matrix. As such, minimizing the trace of the asymptotic (co)variance matrix will minimize the variances (standard errors) of the associated parameter estimates, without consideration being given to the covariances. Given that the trace is calculated as the sum of the diagonal elements, if one of these elements is large in magnitude, then that element will tend to dominate the calculation. For this reason, the $A_{(b)}$-error measure has fallen out of favor. The $D_{(b)}$-error computation is a little more complicated as the determinant of a matrix is calculated as a series of multiplications and subtractions over all the elements of the matrix (see for example, Kanninen 2002). As such, the determinant (and by implication, the $D_{(b)}$-error measure) summarizes all the elements of the matrix in a single 'global' value. Thus, whilst attempts to minimize the D-error measure, on average, minimize all the elements within the matrix, it is possible that in doing so, some elements (variances and/or covariances) may in fact become larger. Despite this property, the $D_{(b)}$-error measure has become the most common measure of statistical efficiency within the literature.

Whatever measure of statistical efficiency is used by the researcher, the generation of an efficient SC design requires that the attribute levels that are assigned to the design be evaluated as to their influence on the asymptotic (co)variance matrix for the appropriate model to be estimated[20] (the second derivatives of the log-likelihood function). The general form of the log-likelihood function for a model of discrete choice can be expressed as equation (6).

$$LL(\beta) = \sum_{n=1}^{N} \sum_{s=1}^{S} \sum_{j} y_{njs} \ln(P_{njs}), \tag{6}$$

where N represents the number of respondents, S the number of choice situations faced by each respondent, j the alternatives present in each s, and y_{njs} a choice indicator taking the value one if alternative j was chosen or zero otherwise. P_{njs} in equation (6) represents the probability that alternative j will be chosen by respondent n in choice situation s.

The presence of P_{njs} in equation (6) plays an extremely important role in generating efficient SC experiments. The probability that an alternative will be selected is a function not only of the attribute levels and priors (parameters) of that alternative, but also of the attribute levels and priors (parameter) assumed for all other competing alternatives. As such, *changing the existing order* of the attribute levels within a design will generally[21] influence the log-likelihood function, and in turn the asymptotic (co)variance matrix for that design. Similarly, the log-likelihood function and asymptotic (co)variance matrix of a design will be equally influenced by the priors assumed in the design generation process. Whereas an orthogonal design will be

[20] This means that the likely model to be estimated (e.g, MNL, NL, ML), be known *a priori* as the derivation of the asymptotic (co)variance matrices of different model forms requires different considerations.

[21] We use the term generally here as the influence on the asymptotic (co)variance matrix is dependent on a large number of factors, not the least of which are the priors assumed for each of the design related parameters.

orthogonal from one experiment to another, an efficient design will generally be efficient only for the specific experiment for which it was created.

The *trick* in generating an efficient SC design is therefore to manipulate the attribute levels of the design and observe the changes in the asymptotic (co)variance matrix given the manipulations made. Unfortunately, even a slight change in one attribute level will likely influence P_{njs} and hence impact upon the entire asymptotic (co)variance matrix for the design. The *unfortunate* part in the above is that the direction and magnitude of the impact will be *largely* unpredictable *a priori* to the change made. It is therefore necessary to manipulate *intelligently* the attribute levels in some way; otherwise the analyst may waste a significant amount of time and computing resources evaluating inefficient design manipulations. In the next section, we discuss methods to reduce the computing time required to locating more efficient SC designs.

4.2 Design challenges

There are a number of significant challenges which face those wishing to generate efficient SC experiments. Aside from the lack of available software capable of generating such designs (only a few programs are currently available including SAS, some GAUSS code and ITLS's NGENE, for example), the level of expertise and the amount of time necessary to generate such designs is currently significantly prohibitive. The level of expertise in generating such designs will naturally improve over time, however, the amount of time required to generate efficient experimental designs will likely remain a problem given greater complexity in the designs that are being generated, even given increases in the computing power available to today's discrete choice modelers. In the following sections, we outline some of the challengers that face those wishing to design efficient SC experiments and what we believe may be some possible solutions to these challengers.

4.2.1 Reducing generation run times

By and far, the greatest problem in generating efficient SC experiments is the amount of time required to generate such designs. A significant number of available software packages (e.g., SAS, SPSS, NGENE, SPEED, CONSURV etc.) are capable of generating orthogonal designs of various dimensions. Mostly, these software packages rely on tables of known orthogonal designs (this is how SPSS orthogonal designs are generated), meaning that where such designs exist, the software package can very quickly generate the desired design. However, fewer software packages are currently available that are capable of generating efficient designs. Where such software packages are available, the generation time is generally far greater than for generating orthogonal designs. Indeed, for all but the smallest of SC experiments, the run time required to locate an efficient design can range anywhere from minutes to even days or weeks, with the amount of time required being a function of the type of econometric models the designs are being generated for as well as, the dimensions of the designs being considered. In the following sections, we discuss some means that are currently being investigated to reduce the computation times required to locate efficient SC designs.

4.2.1.2 Independent random draws versus quasi random

The current literature on the generation of efficient experimental designs using Bayesian methods has tended to rely on independent random *Monte Carlo* draws for priors taken from pre-specified distributions. The results for such methods are highly dependent on the number of draws taken as well as the seed used in generating the draws, a fact that has been well identified and addressed within the mainstream discrete choice modeling literature (Bhat 2001, 2003; Sandor and Train 2003). Typically the literature has tended to use only a small number of draws in an effort to reduce software run times. The use of only a small number (with small being undefined) of independent random draws, however, will likely mean that any efficient design generated will be efficient only for the small number of draws made, and different designs may be deemed efficient given different sets of draws. Even with 1000 or 2000 independent random draws, the average D_b-error for a design given different starting seeds can vary by as much as 10 percent and it is not infeasible that over 100,000 random draws may be required to obtain stability in generating efficient designs.

Rather than rely on independent random draws, several researchers working in other related areas of discrete choice modeling (in particular, on mixed logit models) have examined the use of quasi random draws as a method to reduce the number of draws required to obtain sufficient coverage of distribution space (Bhat 2001, 2003; Sandor and Train 2003). These researchers have shown significant efficiency gains in terms of parameter stability and estimation time when using such methods. Nevertheless, with the exception of Sandor and Wedel (2002), a paper which appears to have been largely ignored within the mainstream experimental design literature, the use of quasi random draws appears to have been overlooked.

Gauss-Hermite approximation

When Bayesian distributions for the priors are assumed normal, stability of selected efficiency measure employed in generating a design, may be achieved by using the Gauss-Hermite approximation method. The approximation works as follows. Let the draws, $B_{k,r}$, r = 1,…, R_k, be designed draws taken from a series of normal distributions, the number of distributions equal to *k*, the number of parameters. Each draw of $B_{k,r}$ is calculated as:

$$B_{k,r} = \mu_k + x_{,k,r} \times \sigma_k \times 2^{0.5} \text{ with associated weights } w_{k,r} / \Pi^{0.5} \quad (7)$$

$x_{,k,r}$ and $w_{k,r} / \Pi^{0.5}$ are taken from the table below, depending on the value of R_k specified by the analyst for each prior B_k. It is necessary to evaluate all combinations of draws such that the total number of evaluations is $R = R_1 \times R_2 \times \ldots R_k$.

Table 9: Gauss-Hermite approximation weights and points

$R_k = 2$	$w_{k,r} / \Pi^{0.5}$	$x_{,k,r}$
0	0.5	-0.70711
1	0.5	0.70711
$R_k = 3$	$w_{k,r} / \Pi^{0.5}$	$x_{,k,r}$
0	0.6667	0
1	0.1667	-1.22474
2	0.1667	1.22474
$R_k = 4$	$w_{k,r} / \Pi^{0.5}$	$x_{,k,r}$
0	0.4541	-0.52465
1	0.4541	0.52465
2	0.0459	-1.65068
3	0.0459	1.65068

Step 1: the analyst determines the numbers R_k (= 2, 3 or 4 corresponding to sheets n=2, n=3 and n =4).

Step 2: create a full factorial for R. For this example, if $R_k = 2$ for each attribute, then the full factorial with have four evaluations (2×2). If $R_k = 3$ for each attribute, then the full factorial will involve nine evaluations (3×3), and if $R_k = 4$ for each attribute, then the full evaluation will involve 16 evaluations (4×4). Note that it is possible to allow a different number for R_k for each attribute (e.g., $R_1 = 2$, $R_2 = 3$, then the total number of evaluations will be six (2×3)). The full factorial is then populated using equation (7), thus providing the full enumeration of R combinations.

Step 3: The total number of draws used in the calculation is equal to R (the full factorial). The R evaluations calculated in *Step 2* are used as the priors in *Step 3*. For each R_k, the efficiency measure (e.g., D_b-error, A_b-error, etc.) is computed as normal.

Step 4: Rather than take the average of the efficiency measure calculated in *Step 3*, the weighting value $w_{k,r} / \Pi^{0.5}$ are applied to each. The correct weights to apply are also calculated from the full factorial. Multiply each efficiency measure value by W and then sum the total. This value is the efficiency measure for the design (equivalent to the average efficiency measure using the Monte Carlo D_b-error but requiring much less draws).

Halton (Sequences) Draws

Halton sequences have been used in the discrete choice literature to provide better coverage of distributional space. Halton sequences are generated in multiple dimensions by selecting an integer, i ($i \geq 2$), and expanding a sequence of integers from one to the desired number of draws, R, using i as the base. The steps in generating Halton sequences are as followed.

Step 1: List the sequence of integers up to R, the total number of draws required (e.g., {0, 1, 2, 3, 4, 5, 6, 7, 8, 9 …, R}).

Step 2: Select $i \geq 2$.

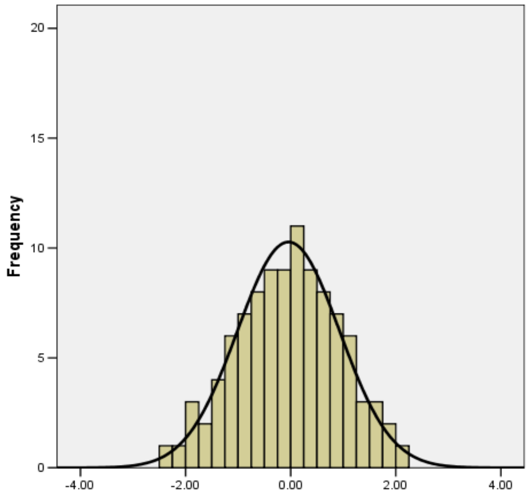
Step 3: Convert the integers to base i selected in step 2. For example, the sequence of integers listed above to base 3 would be {0, 1, 2, 10, 11, 12, 20, 21, 22, 100 … R}. For i = 10, the sequence remains unchanged (i.e., {0, 1, 2, 3, 4, 5, 6, 7, 8, 9 …, R}).

Step 4: Reverse the order of the values for each digit obtained in *Step 3* and reflect the resulting numbers around the decimal point. For the base 3 example shown above, the sequence becomes: {0→0.0, 1→0.1, 2→0.2, 10→ 0.01, 11→0.11, 12→0.12, 20→0.02, 21→0.12, 22→0.22, …, R}. The same sequence in base 10 is: {0→0.0, 1→0.1, 2→0.2, 3→0.3, 4→0.4, 5→0.5, 6→0.6, 7→0.7, 8→0.8, …, 12→0.21, …, R}.

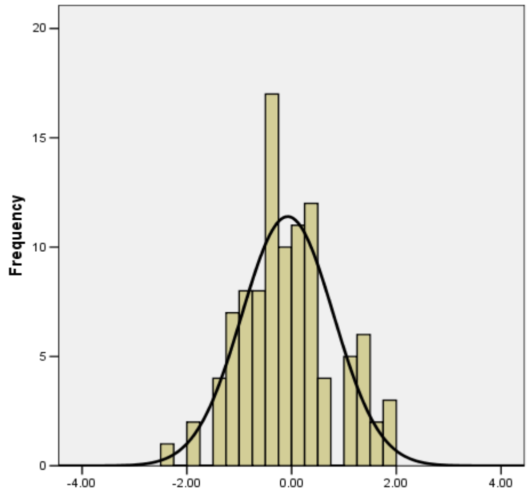
Step 5: Convert the values obtained in *Step 4* back to base 10. For the first sequence (base 3), the Halton sequence is given as: {0.0→0, 0.1→1/3, 0.2→2/3, 0.01→1/9, 0.11→4/9, 0.12→7/9, 0.02→2/9,0.12→5/9, 0.22→8/9, …, R} and the base 10 sequence as: {0.0→0, 0.1→1/2, 0.2→1/4, 0.3→3/4, 0.4→1/8, 0.5→5/8, 0.6→3/8, 0.7→7/8, 0.8→1/16, …, 0.21→3/16, …, R}

Figure 10, shows 100, 250 and 1000 Halton sequence (i = 0) and random draws when applied to a normal distribution with mean zero and standard deviation one. The Halton sequence covers the distributional space much better than independent random draws even at 100 draws, although it should be noted that 1000 random independent draws could feasibly perform better than shown here.

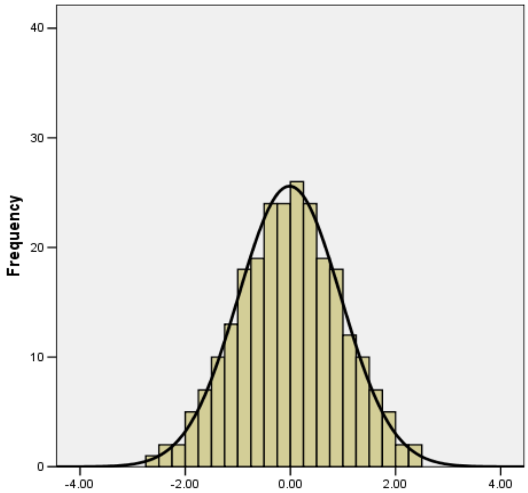
Bliemer and Rose (2006) have explored the use of Gauss-Hermite approximation and Halton sequences in generating Bayesian SC designs. They found that the use of Gauss-Hermite outperforms both independent random and Halton draws for designs with up to eight parameters, but that for designs with greater than eight parameters, Halton draws are preferred.



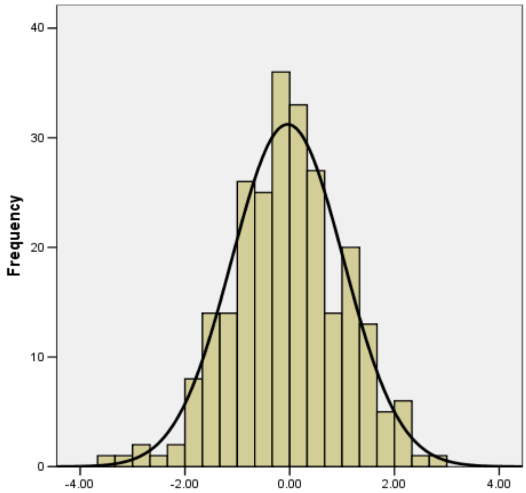
(a) 100 Halton Draws



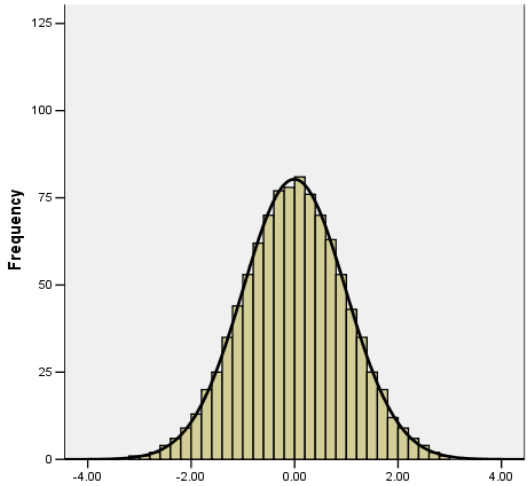
(b) 100 Independent Random Draws



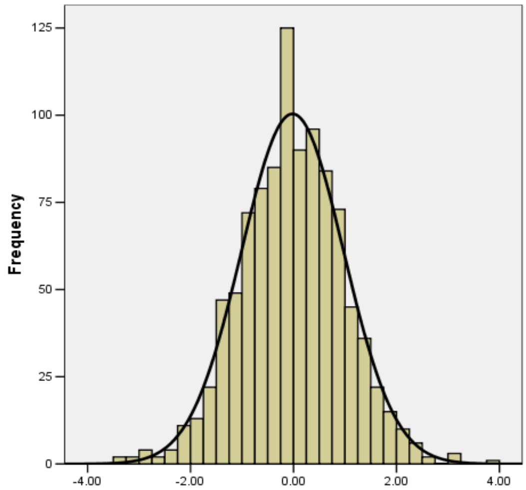
(c) 250 Halton Draws



(d) 250 Independent Random Draws



(e) 1000 Halton Draws



(f) 1000 Independent Random Draws

Fig 10: Halton versus independent random draws assigned to a N(0,1)

The recent interest in generating SC experiments for mixed logit (ML) models has increased the need to research the use of intelligent draws in generation SC designs. Models such as the ML model assume distributions for one or more parameters of the model (each with a mean and standard deviation). Assuming the true mean and standard deviation of the parameter distributions are not known with certainty prior to the generation of the design, then the values for these population moments should also be drawn from Bayesian prior distributions. As such, the generation of ML designs requires not only draws for the random parameters of the models, but draws reflecting uncertainty of the population moments for each of the random parameters as well. Clearly, this requires significant computing resources to achieve; hence there is a need to invest research effort into the effects of using intelligent draws drawn from intelligent draws in designing efficient SC experiments.

4.2.1.3 Using the asymptotic properties of discrete choice models

There exist at least two approaches in generating and evaluating the properties of an efficient design. The first approach involves simulation of a sample of respondents, *N*, after which Monte Carlo simulations can be used to test the efficiency of the design as applied to the simulated sample. This approach requires that the choice response, the y_{njs} vector in equation (6), be generated for each choice situation. In order to do this, for a given design and known parameters (the priors in this instance), the analyst takes a random draw representing the error component of the model and computes the (dis)utility for each alternative. Once the utilities are known to the analyst, y_{njs} is assigned a value of one for the alternative with the highest (dis)utility or zero otherwise. Once the y_{njs} vector has been simulated for the sample, the desired choice model can be estimated for the design. Given a large enough sample, the level of efficiency for various designs can be evaluated. Whilst relatively straightforward to implement, the use of simulated data requires substantial computation time.

Rather than rely on Monte Carlo simulation methods, it is possible to use the asymptotic properties of the discrete choice models to reduce computation time in evaluating numerous SC experiments. For models of discrete choice, the asymptotic (co)variance matrix is equivalent to the second derivatives of the log-likelihood function for the appropriate model form. For the simple MNL model, it can easily be shown that the choice profile, y_{njs}, in equation (6) disappears when the second derivatives are taken (see Bliemer and Rose 2005a). As such, knowledge of the vector of the choice profile, y_{njs}, is not necessary in order to evaluate the asymptotic (co)variance matrix for this model form. Nevertheless, the y_{njs} vector in other models of discrete choice do not disappear when the second derivatives are taken from the log-likelihood function, meaning that knowledge of this vector *is* required in order to evaluate the second derivatives of such models.

Fortunately, we are interested in the asymptotic properties of efficient SC designs. As such, the main point of interest lies in the asymptotically limiting case of $N \rightarrow +\infty$. As argued by Bliemer *et al.* (2006) and Sandor and Wedel (2002), in large samples, the asymptotic properties of discrete choice models allow the substitution of P_{js} (the probability of choosing alternative *j* in choice situation *s*) for y_{njs} given that $P_{js} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} y_{njs}$. Following from this substitution, the sub-index *n* will no longer be present within the second derivatives of the log-likelihood functions, as the summation over the respondents is simply the multiplicand of the value by *N*.

Given the above, it is therefore possible to generate a single design (or rather a design for a single individual) and substitute the y_{js} vector (subscript n = 1, and hence drops out) with the vector of probabilities, P_{js}, which represent the choice probabilities over the entire sample of respondents, *N*. In this way, there is no need to simulate data, or estimate models in order to obtain the asymptotic (co)variance matrix for a design. As shown in Bliemer and Rose (2005a), once a design is generated for a single individual, the asymptotic (co)variance matrix for the design can be simply divided by *N* to establish what it would look like at that sample size (this is why *N* also appears in equations (4) and (5) as $\frac{1}{N}$).

Table 10 demonstrates this result precisely. Part (a) of Table 10 shows the asymptotic (co)variance matrix for a design generated using the probability substitution method described above, as applied to a single respondent. Part (b) of the table shows the results of a Monte Carlo simulation for the same design in which 2500 respondents were simulated using the same parameter priors used in generating the asymptotic (co)variance matrix shown in part (a) of the table. Dividing each element of the asymptotic (co)variance matrix shown in part (a) by 2500 reproduces exactly the asymptotic (co)variance matrix shown in part (b) of the table. This is shown in part (c) of the table. This result will hold for any sample size[22].

For models which offer a closed form solution[23] for the second derivatives of the log-likelihood function (such as the MNL and NL[24] models), the asymptotic (co)variance matrix for a design can be analytically derived. For models with an open form solution when taking the second derivatives of the log-likelihood function (such as the ML and probit models), it is necessary to resort to numerical approximations of the first and second order derivatives of the model. This is usually accomplished using simulation methods[25]. In terms of the generation of efficient SC designs, the need to numerically derive the asymptotic (co)variance matrix for designs associated with open form solutions presents a significant problem for the analyst. Most researchers advocate the use of the DFP and BFGS algorithms (see Train 2003), however, these methods rely on iteratively maximising the log-likelihood function of some data, and updating the

[22] The literature on the generation of efficient SC designs appears to be wedded to the use of Monte Carlo simulations to test the influence of sample size on the efficiency of experimental designs. Given the above results, this appears to preposterous, as one can determine the influence of sample size based on a design generated for a single individual (see Bliemer and Rose 2005b and Rose and Bliemer 2005 for further information on sample size and efficient SC design generation).

[23] Closed form refers to the fact that when taking the (second) derivatives of a function, no integration term remains within in the resulting derivative. Open form refers to models in which the (second) derivatives of a function, an integration term remains. In such cases, the (second) derivative cannot be analytically evaluated for such models.

[24] Whilst it is well known that the NL model offers a closed form solution when deriving the asymptotic (co)variance matrix, an examination of available software capable of estimating such models suggests that numerical approximation of the asymptotic (co)variance matrix is employed in model estimation, as opposed to the analytical derivation of the matrix. For example, Nlogit defaults to the BFGS algorithm to numerically compute the asymptotic (co)variance matrix for NL models. Bliemer *et al.* (2006) derive the analytical equations for the asymptotic (co)variance matrix for NL models with two levels. The use of analytically derived as opposed to numerically approximated asymptotic (co)variance matrices should generally result in quicker model estimation (or in the case of design generation, quicker and more accurate representation of the asymptotic (co)variance matrix of a design).

[25] Numerous algorithms exist for this, including the BHHH, DFP and BFGS algorithms (see Train 2003).

asymptotic (co)variance matrix given the results of the previous iteration. In effect, this necessitates the estimation of the model. The BHHH algorithm, however, does not require that the asymptotic (co)variance matrix be updated over iterations, and as such, possibly offers the best way forward in evaluating efficient SC designs for complex discrete choice models.

Table 10: Asymptotic (co)variance matrix for a design using probability substitution and Monte Carlo simulations

(a) Probability substitution with *N* = 1

	β_1	β_2	β_3	β_4	β_5	β_6	β_7
β_1	**0.1245**	0.0691	0.0084	0.1505	-0.3928	0.0873	0.1636
β_2	0.0691	**0.0748**	0.0188	0.1000	-0.2499	0.0528	0.1166
β_3	0.0084	0.0188	**2.1117**	0.0974	6.5426	0.0215	0.0497
β_4	0.1505	0.1000	0.0974	**0.3391**	0.3136	0.1297	0.2527
β_5	-0.3928	-0.2499	6.5426	0.3136	**33.4178**	-1.0807	-1.1862
β_6	0.0873	0.0528	0.0215	0.1297	-1.0807	**0.2942**	0.1257
β_7	0.1636	0.1166	0.0497	0.2527	-1.1862	0.1257	**0.3932**

(b) Sample generation Monte Carlo method with *N* = 2500

	β_1	β_2	β_3	β_4	β_5	β_6	β_7
β_1	$\mathbf{4.98\times10^{-5}}$	2.76×10^{-5}	3.35×10^{-6}	6.02×10^{-5}	-1.57×10^{-4}	3.49×10^{-5}	6.54×10^{-5}
β_2	2.76×10^{-5}	$\mathbf{2.99\times10^{-5}}$	7.52×10^{-6}	4.00×10^{-5}	-9.99×10^{-5}	2.11×10^{-5}	4.67×10^{-5}
β_3	3.35×10^{-6}	7.52×10^{-6}	$\mathbf{8.45\times10^{-4}}$	3.89×10^{-5}	2.62×10^{-3}	8.59×10^{-6}	1.99×10^{-5}
β_4	6.02×10^{-5}	4.00×10^{-5}	3.89×10^{-5}	$\mathbf{1.36\times10^{-4}}$	1.25×10^{-4}	5.19×10^{-5}	1.01×10^{-4}
β_5	-1.57×10^{-4}	-9.99×10^{-5}	2.62×10^{-3}	1.25×10^{-4}	$\mathbf{1.34\times10^{-2}}$	-4.32×10^{-4}	-4.74×10^{-4}
β_6	3.49×10^{-5}	2.11×10^{-5}	8.59×10^{-6}	5.19×10^{-5}	-4.32×10^{-4}	$\mathbf{1.18\times10^{-4}}$	5.03×10^{-5}
β_7	6.54×10^{-5}	4.67×10^{-5}	1.99×10^{-5}	1.01×10^{-4}	-4.74×10^{-4}	5.03×10^{-5}	$\mathbf{1.57\times10^{-4}}$

(c) Probability substitution with *N* = 2500

	β_1	β_2	β_3	β_4	β_5	β_6	β_7
β_1	$\mathbf{4.98\times10^{-5}}$	2.76×10^{-5}	3.35×10^{-6}	6.02×10^{-5}	-1.57×10^{-4}	3.49×10^{-5}	6.54×10^{-5}
β_2	2.76×10^{-5}	$\mathbf{2.99\times10^{-5}}$	7.52×10^{-6}	4.00×10^{-5}	-9.99×10^{-5}	2.11×10^{-5}	4.67×10^{-5}
β_3	3.35×10^{-6}	7.52×10^{-6}	$\mathbf{8.45\times10^{-4}}$	3.89×10^{-5}	2.62×10^{-3}	8.59×10^{-6}	1.99×10^{-5}
β_4	6.02×10^{-5}	4.00×10^{-5}	3.89×10^{-5}	$\mathbf{1.36\times10^{-4}}$	1.25×10^{-4}	5.19×10^{-5}	1.01×10^{-4}
β_5	-1.57×10^{-4}	-9.99×10^{-5}	2.62×10^{-3}	1.25×10^{-4}	$\mathbf{1.34\times10^{-2}}$	-4.32×10^{-4}	-4.74×10^{-4}
β_6	3.49×10^{-5}	2.11×10^{-5}	8.59×10^{-6}	5.19×10^{-5}	-4.32×10^{-4}	$\mathbf{1.18\times10^{-4}}$	5.03×10^{-5}
β_7	6.54×10^{-5}	4.67×10^{-5}	1.99×10^{-5}	1.01×10^{-4}	-4.74×10^{-4}	5.03×10^{-5}	$\mathbf{1.57\times10^{-4}}$

4.2.1.4 Distributed networks

If, for example, it takes one computer 24 hours to locate an efficient design, then it is feasible that it could take four computers six hours to locate the same design, assuming each computer were able to communicate with the other as to what it is doing so as to avoid repetition of effort. Rather than look towards improvements in computing power, one possible way forward in generating statistically efficient SC experiments is to move towards the use of distributed computer networks. Work currently being conducted at the Institute of Transport and Logistics Studies, Sydney (unpublished) has shown remarkable gains in computation time given the use of distributed computer networks in generating efficient SC designs, however, it should be noted that the use of such networks comes with a barrier of significant up front programming which may prove prohibitive to many researchers.

4.2.1.5 Smarter algorithms

In order to search over the available design space, some form of algorithm is required to manipulate the order of the attribute levels of the design. Within the SC experimental design literature, there appears to exist only a limited number of algorithms in use. Initial studies in the generation of efficient SC designs (Kufheld *et al.* 1994; Zwerina *et al.* 1996) were limited to the Modified Fedorov exchange algorithm. This algorithm, originally developed for generating efficient designs for linear models (Fedorov 1972) begins by generating a pre-specified set of candidate choice situations, some of which are initially assigned to the design to be constructed. The choice situations assigned to the design are systematically exchanged for other choice situations from the candidature set, and retained if an improvement in efficiency is observed to be achieved. The Modified Fedorov exchange algorithm, whilst simple, will generally result in a local efficient design as the total design space explored will be limited to the candidate set generated as part of the algorithm.

Recently, the SC experimental design literature has moved towards the use of an algorithm known as the Relabelling, Swapping and Cycling (RSC) algorithm (or derivations thereof; e.g., the RS algorithm of Huber and Zwerina 1996). Relabelling in the RSC algorithm occurs by exchanging the attribute level values within an attribute with each one another (e.g., for attribute A, 1→3, 2→4, 4→2 and 3→1). If the exchange yields a more efficient design based on whatever criteria is selected by the researcher, then the corresponding design is retained. One benefit of relabelling is that for small designs, it is generally possible to explore each possible permutation over all attributes of the design in a relatively small amount of time. This becomes much more difficult for larger designs, however.

Swapping in the RSC algorithm occurs by simply swapping two attribute levels within a choice situation whilst all other attribute levels remain in place. The algorithm has algorithm has also been implemented using simultaneous swapping of attribute levels (see Kessel *et al.* 2006). The design judged best on the efficiency criteria employed represents the final design to be used as part of the analyst's ongoing study. Cycling of a design is a simple process whereby the attribute levels of the design are exchanged in order, one choice situation at a time, such that 1→2, 2→3, 3→4 and 4→1, etc. This process is continued until the initial design is obtained once more. The best design judged on whatever criteria is then retained.

Combined, the RSC algorithm is generally applied to a design in the order that the name implies. The best design possible is first located using the relabelling algorithm after which the swapping algorithm is employed to determine if yet a better design can be located. Finally cycling is applied to the most efficient RS design. The resulting design should be, at or near optimality (see Kessel *et al*. 2006 and Ferini and Scarpa 2005 for two excellent reviews of the RSC and Modified Fedorov exchange algorithms).

The RSC and Modified Fedorov exchange algorithms have been used extensively in the literature on the generation of SC experimental designs. Unfortunately, the literature on the generation of experimental designs has been limited to small designs (a maximum of eight parameters is the largest design we are aware of and even then this was as a result of the effects coding of the attributes of the design; Kessel *et al*. 2006). Further, apart from Bliemer and Rose (2005a), no single study that we are aware of has properly addressed the issue of alternative specific parameter estimates (Ferini and Scarpa (2005) and Carlsson and Martinsson (2003) come closest by allowing for alternative specific constant terms). Whilst Bliemer and Rose (2005a) use a simple swapping method, the feasibility of the RSC and Modified Fedorov exchange algorithms remains an open question when applied to truly alternative specific (i.e., with parameters other than the constant terms being specified as alternative specific) designs as well as to designs much larger than those currently explored within the literature.

A number of other algorithms are currently under investigation. Bliemer (2006) examines the use of a genetic algorithm with promising results, whilst Collins *et al.* 2006, compare a number of other potential algorithms, also with promising results. One such algorithm, which Collins *et al.* (2006) term *targeted swapping* has been shown to produce impressive results with minimal computation time, even with large designs. This algorithm uses the probabilities of the alternatives to intelligently swap the attribute levels within an attribute. Whilst a perfect utility balanced design (where the utilities are all the same and hence, so to the probabilities of the alternatives; see Huber and Zwerina 1996) may prove too restrictive (as well as extremely difficult to generate, particularly for designs with alternative specific parameter estimates), and may not necessarily result in the best design, Collins *et al.* (2006) note that by swapping the attribute levels of a design in a manner that brings the probabilities closer to balance (but not necessarily perfectly balanced) may under certain circumstances, produce more efficient designs. Rather than naively relabelling, swapping and cycling through the design permutations, Collins *et al.* (2006) found that by intelligently swapping the levels of the design to bring about near utility balance, a design at least as efficient as an efficient RSC design can be located in significantly less time, particularly for designs with many more parameters than eight.

4.3 Working with reference alternatives

In Section 2.3, we briefly introduced the concept of using reference alternatives in SC studies. Indeed, the use of a respondent's knowledge base to derive the attribute levels of the experiment has come about in recognition of a number of supporting theories in behavioural and cognitive psychology, and economics, such as prospect theory, case-based decisions theory and minimum-regret theory (see Starmer 2000; Hensher 2004; Kahnemann and Tversky 1979; Gilboa *et al*. 2002). The use of reference alternatives in SC tasks, however, is inconsistent with current methodology on the generation of efficient SC experiments, highlighting the need to assess SC designs on both statistical and behavioral criteria.

The common use of a fixed set of attribute levels from which to draw from in generating efficient SC experiments is convenient and allows, when priors are assumed, the estimation of the utility functions for the design as well as the related choice probabilities. These in turn may be used to construct the asymptotic (co)variance matrices necessary for determining the efficiency of different experimental designs. However, when the attribute levels of a SC experiment are pivoted as percentages around some base reference alternative, consisting of the attribute levels reported by individual respondents during the survey task, the precise (absolute) attribute levels will not be known to the analyst prior to conducting the survey. As such, the analyst cannot easily determine the statistical efficiency of different designs before going to field. Nevertheless, there exist a number of different strategies that one may use to derive efficient SC experiments, involving the use of pivoting from reference alternatives.

Given the desirability in using reference alternatives in SC experiments, Rose *et al.* (2005) examine a number of possible methods to generate efficient reference based experiments. Strategies examined by Rose *et al.* (2005) include the use of predicted *average attribute levels*, which may be substituted for the fixed attribute levels used in more traditional design generation processes. However, given that designs that rely on the use of reference alternatives employ percentages to *pivot* the attribute levels of the SC alternatives around the fixed alternative specific reference alternative, the use of *average predicted attribute levels* is used simply to determine the percentage differences for the design that will be applied to the real reference alternative for each respondent. That is, once the actual attribute levels of the reference alternative for a respondent are made known to the researcher, the percentage differences are applied to generate the SC alternatives for the study. In using a single population average for the attribute levels, the percentages applied for each choice situation remain fixed over the sampled population (the absolute values differ, however). Rather than use a single 'population average' to predetermine the allocation of the pivot percentages over a design, it is also possible to use segment specific attribute levels (e.g., based on trip length for example) to generate a number of (percentage) designs which are allocated to respondents based on each respondents real 'reference alternative' is determined by the researcher.

Rose *et al.* (2005) also examine the possible use of a two-stage process, whereby information is first captured about the reference alternative in a phase 1 survey, after which either a respondent specific efficient design is generated or a sample specific efficient design is generated after all respondents complete phase one of the project. Once generated, the efficient design can be subsequently administered to the respondent during phase two of the study. In reality, such an approach to SC experiments may prove logistically difficult, however, and statistical efficiency may be lost if not all individuals complete the second phase of the survey.

The use of internet or computer aided personal interviews (CAPI) provide yet another alternative strategy in the generation of statistically efficient SC experiments using reference alternatives. Depending on how the survey is structured, if the information about the reference alternative is captured early in the survey, it may be possible to generate individual-specific efficient SC designs within a single instrument. Nevertheless, we would expect that optimizing designs for each individual would, overall, produce a sub-optimal result in comparison to the proposed two-stage process (assuming zero respondent attrition).

Rose *et al.* (2005) found that orthogonal designs performed relatively poorly when applied to pivot designs, but surprisingly, that when applied to a numerical example, the generation of pre-defined designs based on population and segment-specific reference alternative averages, produced highly reliable parameter estimates, which in some cases were comparable in efficiency to the two-stage and individually optimized designs. Nevertheless, when comparing all asymptotic *t*-ratios of the design methods, the later two strategies do appear to perform best overall. This is to be expected. The objective of producing statistically efficient designs is to minimize the asymptotic standard errors obtained from models estimated from data collected from sampled individuals. Given that the econometric models used for modeling SC data are typically estimated on data pooled from all sampled individuals, it stands to reason that generating a design that minimizes the asymptotic standard errors for the pooled data rather than minimizing the asymptotic standard errors for individuals, we would expect to achieve better results.

Further, we would anticipate that in reality tailoring the design for each sampled individual would produce more efficient designs than the use of assumed averages or even the use of a randomly generated orthogonal design.

5. Improving forecasting accuracy

"APA encourages planners to use reference class forecasting in addition to traditional methods as a way to improve accuracy. The reference class forecasting method is beneficial for non-routine projects such as stadiums, museums, exhibit centers, and other local one-off projects. Planners should never rely solely on civil engineering technology as a way to generate project forecasts" (the American Planning Association 2005).

The promotion of referencing in the micro-behavioral modeling of travel choice in Section 2 and in pivot-designs in Section 4 is linked to the broader theme of forecasting accuracy, as eloquently promoted by Flyvbjerg (2005), who has shown in a number of studies that the average inaccuracy for rail passenger forecasts is -51.4 percent, with 84 percent of all rail projects being in error by more than ±20 percent. For roads, the average inaccuracy in traffic forecasts is 9.5 percent, with half of all road forecasts being in error by more than ±20 percent.

Substantial resources have been spent on improving data and forecasting models with little effect on the accuracy of forecasts (Flyvbjerg, *et al.* 2004). The evidence suggests that something other than poor data and models is at play in generating inaccurate forecasts. Flyvbjerg suggests that psychological explanations account for inaccuracy in terms of optimism bias, that is, '…a cognitive predisposition found with most people to judge events in the future in a more positive light than is warranted by actual experience.' This observation is not new (see Kahneman and Tversky 1979a, Gilboa and Schmeilder 2001 and Starmer 2000), but it is profound in the way that Flyvbjerg crafts the argument in terms of optimism bias[26] and a general failure of traditional approaches that still dominate transport forecasting in the way they embed the relative sophistication of advances in behavioural choice modeling into what might be best described as 'putting the new heart into a very old body'.

Kahneman and Tversky (1979a, b) found human judgment to be generally optimistic due to overconfidence and insufficient regard to distributional information. Thus, people will underestimate the costs, completion times, and risks of planned actions, whereas they will overestimate the benefits of the same actions. Lovallo and Kahneman (2003: 58) call such common behavior the "planning fallacy" and they argue that it stems from actors taking an "inside view" focusing on the constituents of the specific planned action rather than on the outcomes of similar actions that have already been completed.

The common link between the micro and macro perspectives is that errors of judgment are often systematic and predictable rather than random. In particular, we promote the view that drawing on past experiences and evidence within the context of referencing, through the inclusion in choice models of econometric constructs drawn from case

[26] Although not the focus herein, Flyvbjerg (2005) argues that in addition to psychological explanations, political explanations, on the other hand, explain inaccuracy in terms of strategic misrepresentation. "Here, when forecasting the outcomes of projects, forecasters and planners deliberately and strategically overestimate benefits and underestimate costs in order to increase the likelihood that it is their projects, and not the competition's, that gain approval and funding. Strategic misrepresentation can be traced to political and organizational pressures, for instance competition for scarce funds or jockeying for position."

based decision theory and cognitive processing theory, will provide a mechanism for accounting for elements of optimism bias that Flyvbjerg suggests should be used in adjusting forecasts after the inside-view is used to develop forecasts. Our approach recognizes an opportunity to integrate an outside-view within the behavioral specification of the choice modeling system to improve the forecasts of patronage from within. This is equivalent to internalizing the wisdom of the outside view that is on offer from previous projects, except that this wisdom is drawn from deep within the behavioral heart of such project assessments where the real decisions are made on preferences and outcomes. Recent research by the authors is consistent with this direction (Hensher *et al.* 2005a).

6. Conclusions and future directions

Serious efforts are being made to advance the state of econometric tools utilized in the modeling of choice data. The underlying motivation in the development of new statistical techniques is to increase the inferential power available to the analyst, given the predominant methods of both collecting choice data and the general frameworks within which the data are analyzed. That is, researchers seek to minimize the degree to which unobserved effects interfere with the ability of the analyst to make behavioral inference with respect to a given set of choice data.

The inherent limitation of this line of research is that it fails to address sources of misspecification bias that cannot be mitigated without direct methodological approaches. The state of practice tends to abstract from some systematic forces that may influence choice behavior significantly. Failing to incorporate these forces into empirical investigations of choice may lead to misspecification bias that trumps the relative benefits of utilizing advanced statistical techniques. In other words, although it is of merit to advance our statistical toolkit in efforts to account for forces such as preference heterogeneity, it may be of relatively greater merit to seek data collection techniques and general modeling structures beyond the scope of the status quo, in an effort to internalize elements influencing choice behavior that have been generally abstracted from to this point.

This paper promotes two areas in which research effort would be particularly well placed: attribute processing strategies of respondents, and interdependent choice behavior. The predominant assumption that all decision makers attend to all information presented to them equally when making all decisions has been violated in empirical studies of the APSs utilized by respondents. Heuristic decision-making theories proposed by cognitive psychologists and behavioral economists have been supported by observed choices, in which respondents indicate, sometimes overwhelmingly, that rational coping strategies were enacted to attend to a subset of the information presented when making choices. The divergence in behavioral implications across models incorporating APSs versus those that do not can be staggering. Hence, it is clear that responsible studies of choice behavior cannot reply on assumptions of passive bounded rationality, and should take appropriate steps to internalize APS heterogeneity.

Similarly, as supported by McFadden (2001), there is a need to develop empirical strategies to incorporate the effects of interdependency amongst decision makers. Treating all decision makers as though they act in a vacuum is simply unrealistic, and is likely to lead to erroneous inference with respect to choice behavior in settings involving interactivity amongst decision makers. We acknowledge that, in many applications involving choices within groups, it is implausible to collect choice data from all group members simultaneously. However, it is possible to extend choice experiments given to one respondent at a time, to include interdependent elements in both the specification of the choice setting, and the choices made themselves. The techniques highlighted in this paper demonstrate this capacity, and encourage further innovative developments on this front.

In Section 3, we discussed choice studies involving multiple agents making decisions requiring joint co-operation in order to derive beneficial outcomes commiserate to the level of power related to the negotiating parties. Depending on the modelling strategy to be adopted by the researcher, the generation of an efficient SC experiment for such problems may simply represent an extension of current the design methods discussed in Section 4. In the simplest case, design problems involving multiple decision makers may be viewed as nothing more than the combining of two different data sources, similar to SP/RP applications. When treated in this way, the design of the SC experiment will be an adoption of efficient SC experiments for nested logit models (Bliemer *et al.* 2006). Problems leading to a loss of efficiency of the design may occur however when the survey approach follows a game theoretical structure and this alone suggests the need for further research inquiry, which is active in ITLS. Two other features of the design of choice experiments that need more consideration are the inclusion of attribute-processing strategies which may modify the established statistical properties prior to model estimation and designs that include covariates.

Finally, we are currently unaware of any study that has looked at a comparison of efficient versus orthogonal SC designs when applied to real respondents. Current research appears to rely solely on the use of Monte Carlo simulation to predict the efficiency gains obtained in using efficient SC designs. There is plenty more to do despite some notable progress to date.

References

American Planning Association (2005) JAPA Article Calls on Planners to Help End Inaccuracies in Public Project Revenue Forecasting

http://www.planning.org/newsreleases/2005/ftp040705.htm, April 7.

Bhat, C.R., (2001) Quasi-random maximum simulated likelihood estimation of the mixed logit model, *Transportation Research B*, 35(7), 677-693.

Bhat, C.R., (2003) Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences, *Transportation Research B*, 37(9), 837-855.

Berg, N. (2005) Decision-making environments in which unboundedly rational decision makers choose to ignore relevant information, *Global Business and Economics Review*, 7 (1), 59-73.

Bliemer, M.C.J. (2006) Genetic Algorithms for generating optimal experimental designs, working paper, Delft University.

Bliemer, M.C.J. and Rose, J.M. (2005a) Efficient Designs for Alternative Specific Choice Experiments, University of Sydney, February.

Bliemer, M.C.J. and Rose, J.M. (2005b) Efficiency and Sample Size Requirements for Stated Choice Studies, Institute of Transport and Logistics Studies, University of Sydney, (submitted to *Transportation Research B*, August 2005).

Bliemer, M.C.J., Rose, J.M. (2006) Quasi-random Bayesian Efficient Stated Choice Designs, Institute of Transport and Logistics Studies, University of Sydney, February.

Bliemer, M.C.J., Rose, J.M. and Hensher, D.A. (*forthcoming*) Constructing efficient stated choice experiments allowing for differences in error variances across subsets of alternatives, accepted *Transportation Research B*, December 2005.

Brown, M.B. and Forsythe, A.B. (1974) Robust tests for equality of variances, *Journal of American Statistical Association,* 69, 264-267.

Bunch, D.S., Louviere, J.J. and Anderson, D.A (1996) A Comparison of Experimental Design Strategies for Choice-Based Conjoint Analysis with Generic-Attribute Multinomial Logit Models, Working Paper, Graduate School of Management, University of California, Davis.

Carlsson, F. and Martinsson, P. (2003) Design techniques for stated preference methods in health economics, *Health Economics*, 12, 281-294.

Collins, A.T. Bliemer, M.C.J. and Rose, J.M. (2006) A comparision of algorithms for designing efficient stated choice experiments, Institute of Transport and Logistics Studies, University of Sydney, February.

DePalma, A., Meyers, G.M. and Papageorgiou, Y.Y. (1994) Rational choice under an imperfect ability to choose, *American Economic Review*, 84, 419-440.

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.

DeShazo, J.R. and Fermo, G. (2004) Implications of rationally-adaptive pre-choice behavior for the design and estimation of choice models, working paper, School of Public Policy and Social Research, University of Califormia at Los Angeles, August.

Diederich, A. (2003) MDFT account of decision making under time pressure. *Psychonomic Bulletin & Review, 10*, 157-166.

Dosman, D. and Adamowicz, W. (2003) Combining stated and revealed preference data to construct an empirical examination of intrahousehold bargaining, Department of Rural Economy, University of Alberta, December.

Drolet, A. & Luce, M.F. (2004) The rationalizing effects of cognitive load on emotion-based trade-off avoidance. *Journal of Consumer Research, 31*, 63-77.

Erber, R. and Fiske, S.T. (1984) Outcome dependency and attention to inconsistent information, *Journal of Personality and Social Psychology*, 47 (4), 709-26.

Fedorov, V. V. (1972) *Theory of Optimal Experiments*, New York: Academic Press.

Ferini, S. and Scarpa, R. (2005) Experimental designs for environmental Evaluation with choice-experiments: a Monte Carlo investigation, working paper, Waikato Management School, Hamilton, New Zealand.

Flyvbjerg, B. (2005) Beyond the Planning Fallacy: Reference Class Forecasting in Practice, Dept. of Development and Planning, Aalborg University, Denmark.

Flyvbjerg, B., Holm, M. and Buhl, S. (2004) What causes cost overrun in transport infrastructure projects? *Transport Reviews*, 24 (1), 3-18.

Gilboa, I. and Schmeidler, D. (2001) *A Theory of Case-Based Decisions*, Cambridge University Press, Cambridge.

Gilboa, I., Schmeidler, D. and Wakker, P. (2002) Utility in case-based decision theory, *Journal of Economic Theory*, 105, 483-502.

González-Vallejo, C. (2002). Making trade-offs: a probabilistic and context-sensitive model of choice behavior. *Psychological Review, 109*, 137-155.

Greene, W.H., Hensher, D.A. and Rose, J.M (2006) Accounting for heterogeneity in the variance of unobserved effects in mixed logit models, *Transportation Research B.*

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. and Barnard, P.O. (1990) The Orthogonality issue in stated choice designs, in Fischer, M. Nijkamp, P. and Papageorgiou, Y. (Eds.) in Spatial Choices and Processes, North-Holland, Amsterdam; 265-278.

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

Hensher, D.A. (in press- a) Revealing differences in behavioral response due to the dimensionality of stated choice designs: an initial assessment, *Environment and Resource Economics*.

Hensher, D.A. (2006 - b) How do Respondents Handle Stated Choice Experiments? - Attribute processing strategies under varying information load, *Journal of Applied Econometrics,* 21, 861-878.

Hensher, D.A. and Knowles, L. (2005) Spatial alliances of public transit operators: establishing operator preferences for area management contracts with government, paper presented at the *9th International Lisbon, Portugal*, September.

Hensher, D.A., Rose, J.M and Bertoia, T. (in press) The implications on willingness to pay of a stochastic treatment of attribute processing in stated choice studies, *Transportation Research* E.

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

Hensher, D.A. and Puckett, S.M. (2005) Theoretical and conceptual frameworks for studying agent interaction and choice revelation, Institute of Transport and Logistics Studies, The University of Sydney.

Hensher, D.A., Puckett, S.M. and Rose, J.M (2005a) Agency decision making in freight distribution chains: revealing a parsimonious empirical strategy from alternative behavioural structures, Institute of Transport and Logistics Studies, the University of Sydney, September.

Hensher, D.A. and Puckett, S.M., (2006) Power, Cooperation and Concession in Group Choice Making, Institute of Transport and Logistics Studies, the University of Sydney, September.

Huber, J. and Zwerina, K. (1996) The Importance of utility Balance and Efficient Choice Designs, *Journal of Marketing Research*, 33 (3), 307-317.

Kahnemann, D. and Tversky, A. (1979a) Prospect theory: an analysis of decisions under risk, *Econometrica*, 47 (2), 263-91.

Kahneman, D. and Tversky, A. (1979b) Intuitive prediction: biases and corrective procedures. In S. Makridakis and S. C. Wheelwright, Eds., *Studies in the Management Sciences: Forecasting*, 12 (Amsterdam: North Holland).

Kanninen, B. J. (2002) Optimal Design for Multinomial Choice Experiments, *Journal of Marketing Research*, 39 (2), 214-217.

Kessel, R., Goos, P. and Vandebroek, M. (*forthcoming*) Comparing algorithms and criteria for designing Bayesian conjoint choice experiments, *Journal of Marketing Research*.

Kuhfeld, W.F., Tobias, R.D. and Garratt, M. (1994) Efficient Experimental Design with Marketing Research Applications, *Journal of Marketing Research*, 21 (4), 545-557.

Lazari, A.G. and Anderson, D.A. (1994) Designs of Discrete Choice Experiments for Estimating Both Attribute and Availability Cross Effects, *Journal of Marketing Research*, 31 (3), 375-383.

Loomes, G. and Sugden, R. (1987) Some implications of a more general form of regret theory, *Journal of Economic Theory*, 41 (2), 270-87.

Louviere, J.J. and Timmermans, H.J.P. (1990) Hierarchical information integration applied to residential choice behavior, *Geographical Analysis*, 22, 127–145.

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

Lovallo, D. and Kahneman, D. (2003) Delusions of success: how optimism undermines executives' decisions, *Harvard Business Review*, July, pp 56-63

McFadden, D. (2001) Economic choices: economic decisions of individuals, a set of notes prepared for a lecture at the University of California, Berkeley, 18 March.

Oppewal, H., Louviere, J.J., and Timmermans, H.J.P. (1994) Modeling hierarchical information integration processes with integrated conjoint choice experiments, *Journal of Marketing Research,* 31, 92–105.

Payne, J.W. Bettman, J.R. & Johnson, E.J. (1988) Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534-552.

Payne, J.W., Bettman, J.R., Coupey, E. & Johnson, E.J. (1992) A constructive process view of decision making: Multiple strategies in judgment and choice. *Acta Psychologica, 80*, 107-141.

Rose, J.M., Bliemer, M.C.J., Hensher, D.A. and Collins, A.T. (2005) Designing Efficient Stated Choice Experiments Involving Respondent Based Reference Alternatives, University of Sydney, (submitted to *Transportation Research B*, May 2005).

Sælensminde, K. (1994) The Impact of Choice Inconsistencies in Stated Choice Studies, *Environmental & Resource Economics*, 23 (4), 403-420.

Sandor, Zsolt and Wedel, M. (2001) Designing Conjoint Choice Experiments Using Managers' Prior Beliefs, *Journal of Marketing Research*, 38 (4), 430-444.

Sandor, Z. and Wedel, M. (2002) Profile Construction in Experimental Choice Designs for Mixed Logit Models, *Marketing Science*, 21(4), 455–475.

Sandor, Z. and Train, K.E. (2003) Quasi-random simulation of discrete choice models, *Transportation Research B*, 38(4), 313-327.

Slovic, P. (1995) The Construction of Preference. *American Psychologist, 50*, 364-371.

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.

Stroop, J.R. (1935) Studies of interference in serial verbal reactions, *Journal of Experimental Psychology*, 18, 643-62.

Svenson, O. (1992) Differentiation and Consolidation Theory of human decision making: A frame of reference for the study of pre-and post-decision processes, *Acta Psychologica, 80*, 143-168.

Swait, J. and Adamowicz, W. (2001) The influence of task complexity on consumer choice: A latent class model of decision strategy switching, *Journal of Consumer Research, 28*, 135-148.

Thurstone, L.L. (1927) A law of comparative judgment, *Psychological Review, 23,* 273–286.

Train, K.E. (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge.

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

Wang, D., Jiuqun, L. and Timmermans, H.J.P. (2001) Reducing respondent burden, information processing and incomprehensibility in stated preference surveys: principles and properties of paired conjoint analysis, *Transportation Research Record*, 1768, 71–78.

Zwerina, K., Huber, J. and Kuhfeld, W. F. (1996) A general method for constructing efficient choice designs, working paper, Fuqua School of Business, Duke University, Durham, NC 27708. Available from SAS Institute (http://ftp.sas.com/techsup/download/technote/ts629.pdf).