WORKING PAPER
ITS-WP-99-6

Conjoint Preference Elicitation Methods in the Broader Context of Random Utility Theory Preference Elicitation Methods

by

David A. Hensher,
Jordan J. Louviere &
Joffre Swait

March, 1999

ISSN 1440-3501

Established and supported under the Australian Research Council's Key Centre Program.

INSTITUTE OF TRANSPORT STUDIES
The Australian Key Centre in Transport Management

The University of Sydney and Monash University

ITS

INSTITUTE OF TRANSPORT STUDIES

The Australian Key Centre
in Transport Management

The University of Sydney
and Monash University

Introduction

The purpose of this paper is to place conjoint analysis techniques within the broader framework of preference elicitation techniques that are consistent with the Random Utility Theory (RUT) paradigm. This allows us to accomplish the following objectives:

1. explain how random utility theory provides a level playing field on which to compare preference elicitation methods, and why virtually all conjoint methods can be treated as a special case of a much broader theoretical framework. We achieve this by:

    - discussing wider issues in modelling preferences in the RUT paradigm, the implications for understanding consumer decision processes and practical prediction, and how conjoint analysis methods fit into the bigger picture.

    - discussing how a level playing field allows meaningful comparisons of a variety of preference elicitation methods and sources of preference data (conjoint methods are only one of many types), which in turn allows us to unify many disparate research streams;

    - discussing how a level playing field allows sources of preference data from various elicitation methods to be combined, including the important case of relating sources of preference elicitation data to actual market behaviour;

    - discussing the pros and cons of relaxing the simple error assumptions in basic choice models, and how these allow one to capture individual differences without needing individual-level effects;

    - using three cases studies to illustrate the themes of the paper.

Random Utility Theory is not new; Thurstone (1927) proposed it as a way to understand and model choices between pairs of stimuli. RUT languished until McFadden (1974) provided key theoretical and econometric insights necessary to extend the paradigm to the general case of multiple choices and comparisons (Rushton, 1969 independently developed a similar but less far-reaching approach to modelling revealed choices). Since then, RUT has been applied to a wide range of cases of human judgment, decision-making and choice behaviour, and now represents a general framework for understanding and modelling many types of human behaviour.

Thus, we adopt the view that almost all conjoint analysis techniques can be viewed as a special case of the more general RUT paradigm. Historically conjoint techniques have played an important role in understanding preference formation in marketing and other social sciences, so a contribution can be made by introducing and discussing a more general framework and the role of conjoint analysis methods within it. In fact, it could be argued that the term "conjoint analysis" has passed its "use-by" date, and should be replaced with more specific terms like "Random Utility Choice Modelling" to describe various ways to model preference and choices. This could help to counter the misconception that there is one unique technique called "conjoint," when, in fact, there are many forms and flavours of "conjoint analysis," each of which requires different assumptions and analytical techniques. For example, some conjoint theory and methods are consistent with economic theory but some are not, some methods permit one to combine experimental with actual marketplace choice data but many do not, etc.

Indeed, there are many ways to understand and model preferences and choices, some of which bear scant relation to one another and others that are incompatible, both theoretically and analytically. RUT offers a way to unify many seemingly disparate approaches to understand and model preference formation and choice. Figure 1 represents an overview of the general problems covered by RUT, which can assist our understanding of the role of conjoint analysis methods within RUT. Figure 1 should be regarded as a pedagogical vehicle to help explain why a more general view is required; it is not a theory per se.

Figure 1: A General Framework for Understanding Decision Sequences

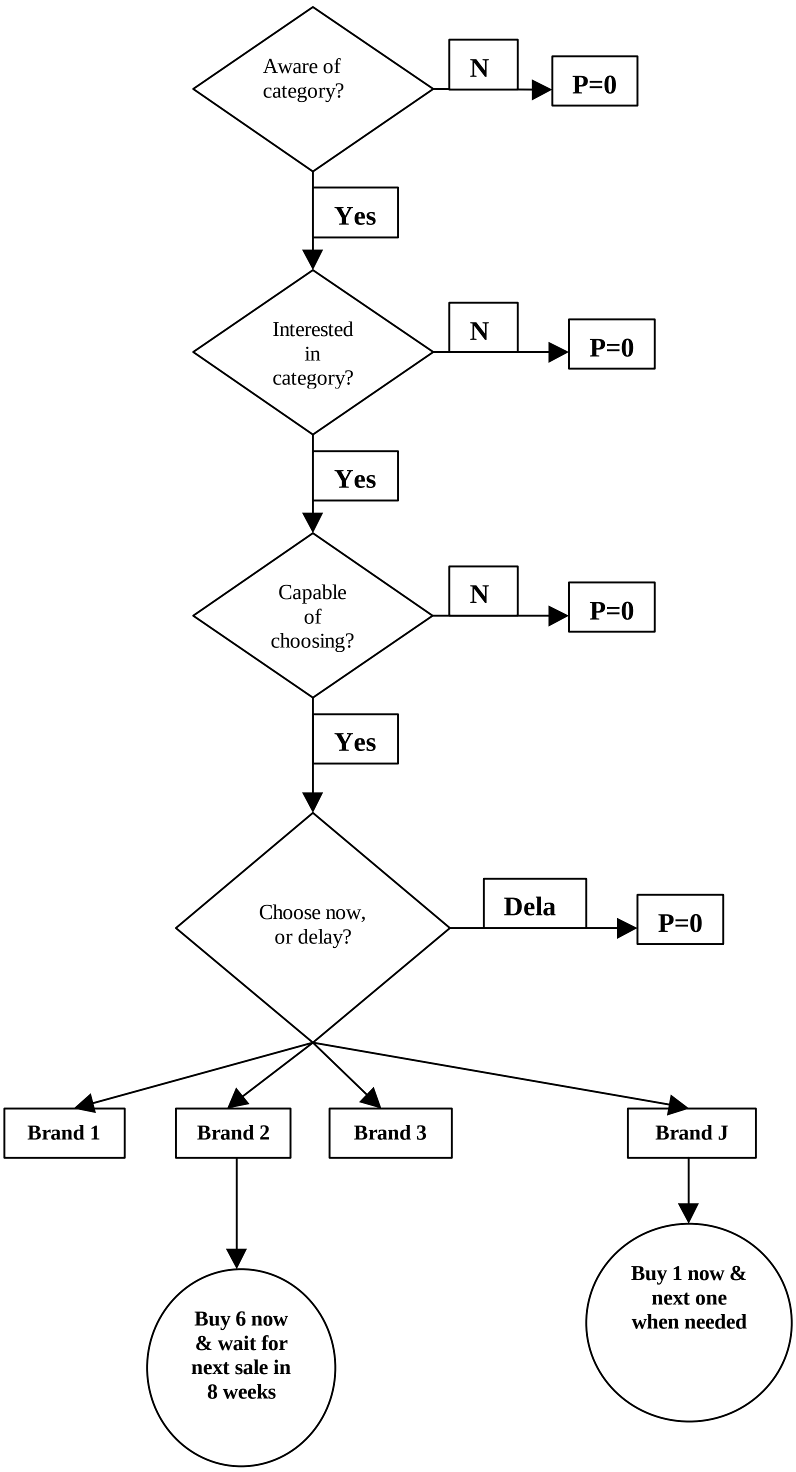


Figure 1 suggests that many consumers make a series of sequential decisions en route to choosing products/services and brands. That is, a consumer first becomes aware that particular product or service categories are available and can satisfy needs/solve problems; if she is unaware of a category, the probability of purchasing brands in it is zero. Once aware, she evaluates benefits or problem solutions offered by category brands to determine her level of interest vis-à-vis purchase costs. If she is not interested in, does not value or cannot use the benefits (eg, she's allergic to certain ingredients) or perceives low value relative to cost, her category brand purchase probabilities will be zero. If she is interested in or perceives value, she then decides if she can purchase. For example, she may need a car, be interested in a luxury car but not be able to afford one or be put off by what her friends will say if she buys one. Thus, if not capable, her category choice probabilities may not be zero, but her brand choice probabilities in the luxury category will be zero. Finally, if aware, interested and capable, she must decide whether to buy now or wait. If she decides to wait, her current period brand purchase probabilities in the category will be zero. If she decides to buy now, she must decide which brand; and in many categories, she can buy more than one brand and more than one quantity of each brand. Thus, in general volume choices depend on brand choices, and inter-purchase times depend not only on brand and volume choices but also on a consumer's personal/household circumstances (eg, income, lifestyle, storage space, etc.).

Although simplistic, the framework allows some interesting and important insights into choice processes and the role of conjoint methods. In particular, many applications of traditional conjoint techniques deal with the evaluation of brands or "brand descriptions" (profiles), while others deal with generic category descriptions, or "concepts" (ie, benefits, problem solutions and costs that profile a category rather than a brand in that category). In the brand case, traditional conjoint methods model consumer evaluations of brand profiles near the end of the decision process sequence, which provides few insights about causally prior processes. Causally prior decision processes can play major roles in market choices, hence are of considerable strategic interest in their own right. In contrast, applications that involve category profile (concept) evaluations are causally prior to brand evaluations, so offer few insights into specific brand, volume or inter-purchase time choices. In general, therefore, they should have relatively low predictive accuracy forecasting trial and repeat choices, except possibly in quite mature categories.

To properly understand and characterise markets, one needs a comprehensive framework to explain and model the types of hierarchical (and temporal) sequences depicted by Figure 1. RUT is such a framework because it recognises that decisions at any stages in a sequence (or any other sequences, including simultaneous decisions not involving sequences) are random utility processes; hence decisions at advanced stages are conditional on decisions at prior stages. Figure 1 suggests why few researchers in marketing report significant individual differences associated with social, demographic, psychographic or similar factors, despite considerable evidence elsewhere that such differences often are significant, particularly in transport applications (eg, Ben-Akiva and Lerman 1985; McFadden 1986; Ben-Akiva and Morikawa 1990). That is, individual differences, such as social and demographic factors, probably play key roles in early stages of decision-making, and explain fewer differences in choices in later stages.

For example, category and brand awareness may be influenced significantly by lifestyle, location of residence, media access and overall social and economic status (inter alia). As well, interest in and capability of purchasing in categories and decisions to delay purchases are likely also to be influenced by such factors. Failure to take such decision sequences and conditioning into account may lead to a wide variety of incorrect inferences and conclusions about decision and choice processes. For example, overall levels of price in categories may play causally prior roles in choices at earlier stages, such that those who choose brands are a "selected" sample of the market. Thus, it is unclear how to interpret price effects in many traditional conjoint or scanner data choice models because such effects are biased if one does not account for decisions not to purchase or delay purchase because of overall category price levels. Consumer income and other social and demographic factors may play roles in such causally prior processes, but are typically ignored, or at least under-represented, in many conjoint and choice modelling applications in marketing. Thus, it is hard to predict market behaviour and future outcomes well if one fails to take causally prior decisions into account when they matter. That is, RUT is not to blame for less than expected prediction accuracy; rather it is the application itself.

A more pragmatic assessment would suggest that decision processes in many markets are relatively stable because these markets are mature; hence, one can ignore prior processes if one's primary objective is short-run prediction. Although the latter observation often may be true, it begs the more general question, which is the object of this paper: ***How can we advance understanding of decision processes and develop better and more accurate approximations to them that will permit us to forecast market behaviour more accurately in both short and long terms?***

The Random Utility Theory Paradigm

The introduction provided insights into deficiencies in academic and applied research in decision and choice processes, including conjoint analysis methods. The purpose of this section is to introduce RUT and explain how it can enhance our understanding of decision processes, compare and contrast models from different sources of preference information, and relate models estimated from sources of stated preference data (eg, conjoint data) to market choices (revealed preferences).

Preference, choice, or more generally, <u>dominance</u> data, come in many forms, such as:

- cross-sections of past, present or future preferences or choices (eg, consumer's last or next brand choices);

- preferences or choices expressed in controlled experiments that manipulate attribute levels and/or construct choice sets (eg, conjoint experiments);

- panels that supply temporal observations of preferences or choices (eg, scanner panels);

- cross-sectional or panel observations of consumer judgments (more generally, "evaluations") of products on latent dimensions like "attractiveness," "intent to purchase," etc. (eg, consumer brand evaluations on magnitude estimation or production scales, rating scales or rank orderings, or "forced" discrete choices involving selection of one or perhaps none of the products);

- direct observations of choices made by single persons or groups of people (eg, direct observation of purchases made from supermarket shelves);

- and many, many more.

This abbreviated list suggests that there are very many possible combinations of preference data types and decision contexts.

To identify relationships among such preference data sources and test for regularity and order, one needs a unified way to compare as many types of preference "revelations" as possible on a level playing field. There has been little progress made in the comparison of such data sources, much less development of mechanisms that can explain differences in them. Indeed, there may be substantial regularities in preference and choice processes, but the lack of a unified framework with which to study and compare them has given rise to a veritable cottage industry of different techniques, measurement methods, experimental paradigms and "stories" used to study and explain decision processes (Meyer, et al. 1999). As Meyer, et al. (1999) explain, there has been limited substantive progress in understanding the processes themselves or regularities that may exist.

As in most scientific endeavours, little progress can be expected without fundamental theory, and empirical comparisons often are not meaningful without it. RUT provides a unified theoretical framework and a theoretically sound and relatively simple way to compare and contrast models estimated from many sources of preference data and develop models to account for real market choices (McFadden 1974, 1981,1986; Ben-Akiva and Lerman 1985; Louviere, Hensher and Swait 1999). In particular, RUT posits that the *"utility"* (or attractiveness) of product/service options can be decomposed into systematic (ie, observed) and random (ie, unobserved) components:

$$U_i = V_i + \varepsilon_i, \quad (1)$$

where U_i is a latent measure of the attractiveness of option i, V_i is an observable, systematic or "explainable" component of the attractiveness of option i, and ε_i, is a random or "unexplainable" component of option i.

Randomness arises because researchers, scientists or analysts cannot "look" into consumers' heads and observe the true attractiveness of each alternative. Instead, they can indirectly observe indicators of the true attractiveness by designing elicitation procedures ("preference elicitation procedures" or PEPs) to give insights into consumer preferences. Regardless of the time and effort one devotes to understanding preferences, some aspects of consumer preferences cannot be unexplained because all factors that drive preferences cannot be identified, unreliability is inherent in the measurement procedures one uses and preferences may vary at different times or situations for the same or different consumers.

Thus, viewed from the perspective of a scientist trying to explain consumers' decision-making processes and/or preferences, consumer preferences must be stochastic even if consumers themselves are perfectly deterministic. Consequently, we seek to model the probability that a randomly chosen consumer will do something like choose a brand, tick a box, report a choice frequency, etc. That is,

$$P(i|A) = P[\ (V_i + \varepsilon_i) > \ldots > (V_j + \varepsilon_j) > \ldots > (V_J + \varepsilon_J)], \text{ for all j in A}, \quad (2)$$

where all terms are as previously defined, except for P(i|A), which is the probability that a consumer chooses action i from the set of all possible actions {A}, from which she might choose. RUT-based choice models are derived by making assumptions about distributions of the random component (ε_i) and deducing the consequences for equation 2. The effects of interest in the deduced model form are captured by specifying V_i to be a function of a set of observables. These observables can be attributes/levels systematically manipulated in conjoint or conjoint-like experiments and/or measures of quantities one hypothesises to explain or drive preferences in particular situations. Regardless of whether the effects are controlled, uncontrolled or some combination of both, V_i typically is expressed as a linear-in-the-parameters function of the observables:

$$V_i = \sum_k \beta_k X_{ki}, \qquad (3)$$

where β is a K-element vector of parameters and X is an i by k vector (matrix) of observables that describe the actions that were available to be chosen and the consumers who make the choices (other relevant factors also can be included, like different situations, conditions, time periods, environments, cultures, etc.). RUT assumes that consumers try to choose those actions (options) that are the most attractive, subject to constraints like time, money, and peer pressure. We are unaware of other widely applied, scientifically accepted and empirically verified choice rules except the maximum utility rule, although others have been proposed from time-to-time (eg, Simon 1983).

The foregoing discussion is well-known to many social scientists, and has been the basis for a vast amount of previous research in the RUT paradigm. However, what is less well-known and less-well appreciated is the fact that the random component itself plays a fundamental role in the behavioural outcomes in many sources of preference data. That is, random components play key roles in the statistical inferences made about model estimates (β's, or so-called "partworths") within data sources, as well as in comparisons of model estimates across data sources (Morikawa 1989; Swait and Louviere 1993; Louviere, Hensher and Swait 1998, 1999).

In particular, the random component (*specifically, the variability in the random component*) is inherently linked to estimates of the partworths, and cannot be separately identified in any one source of preference data (Ben-Akiva and Lerman 1985; Swait and Louviere 1993; Louviere, Hensher and Swait 1998, 1999). As explained by Ben-Akiva and Morikawa (1990) and Swait and Louviere (1993), the variance of the random component is inversely proportional to a constant that "scales" the β parameters in all RUT choice models. In fact, the systematic component is actually $\lambda\beta_k X_{ki}$, where λ is a multiplicative constant that is inversely proportional to the variability of the random component. A scalar λ is embedded in all choice models, regardless of the distributional assumption made about the ε's to derive any particular choice model specification. Thus, λ cannot be separately identified (estimated) in any one source of preference data. However, as explained by Ben-Akiva and Morikawa (1990) and Swait and Louviere (1993), ratios of λ's can be identified from two or more sources of preference data if one source of data is used as a constant reference. We note now, but leave for subsequent discussion, that identification of λ within a single data set is possible if exogenous information is introduced to allow for individual or alternative-specific differences or if variances can be freed up to identification limits (eg, HEV models). We note that the present discussion assumes that the random components (ie, error terms) are independent and identically distributed.

These "variance-scale ratios" play crucial roles in comparing models estimated from different sources of preference data and testing hypotheses about how and why they might differ. That is, one must account for differences in random component variability to compare different sources of preference and choice data and rule out the possibility that data sources differ only in levels of error variability (ie, size of random components) before concluding that differences in model parameters are real. Swait and Louviere (1993) show how to test model and process hypotheses by controlling for and taking into account differences in random components.

The relevance of the preceding discussion to the objectives of this paper is that RUT provides the basis for level playing field comparisons of many sources of preference and choice data. As Louviere (1994) noted, data from any conjoint experiment can be couched in a RUT framework, and modelled as a nested sub-model within the RUT family of choice models. This allows one to test if utility model parameters differ after accounting for differences in levels of error variability between tasks, task conditions, time periods, situations, etc., regardless of whether the observed data were generated from rating, ranking, binary, multiple choice or other types of judgment and choice tasks. The basis for this conclusion is the Luce and Suppes (1965) Ranking Theorem, which proves that any form of dominance data can be expressed as some type of weak or strong order, which are consistent with RUT and can be used to estimate RUT-based choice models. However, different sources of preference data generally will have different levels of random error variability, and hence different values of λ. As earlier noted, differences in λ must be taken into account to compare models estimated from different sources of preference data, and failure to do so can lead to misleading conclusions about differences in data sources, decision processes, etc.

The foregoing is now well-known in econometrics, transport and environmental and resource economics, but has been largely ignored in academic and applied marketing research. For example, Ben-Akiva, et al. (1994) discuss and Louviere, Fox and Moore (1993) show how to extend Louviere's (1994) argument to many forms of preference data, which in turn allows preference (utility) parameters to be compared and tested for different sources of preference data, types of decision tasks, task manipulations (eg, attribute or profile orderings), groups of people, time periods, etc. Meyer, et al. (1999) review a number of papers in this paradigm, such as Hensher and Bradley (1993); Swait and Louviere (1993); Swait, Louviere and Williams (1994), Adamowicz, et al. (1994, 1996); Swait, Adamowicz and Louviere (1998); Louviere, Hensher and Swait (1998) (to name only a few). As well, many of the basic issues are discussed in Carson, et al (1994); Keane (1997); and Louviere, Hensher and Swait (1998, 1999). Thus, the theory is well-established, there have been numerous empirical tests of the basic ideas discussed above and there is considerable empirical support for the general approach.

Moreover, comparisons of utility model parameters from different preference data sources have been the subject of considerable research attention for many years. A few examples include Meyer (1977), who examined profile order on attribute weights; Meyer and Eagle (1982), who studied the effect of attribute range on model parameters; Johnson (1989) who investigated attribute order effects; Olsen and Swait (1995) and Olsen, et al. (1995), who studied the effects on model parameters of differences in response mode, sequential vs simultaneous presentation, prior practice, etc; Oliphant, et al. (1992), who compared model parameter differences due to ratings and choices and profile or choice set order within tasks; Elrod, et al. (1993), who investigated the effects of Pareto Optimal sets on ratings and choices; Chrzan (1994), who studied order effects in choice tasks; and Ben-Akiva, et al. (1991) and Bradley and Daly (1994), who compared MNL model parameters estimated from different depths of preference ranking data (to name only a few papers). These references demonstrate that RUT now makes it possible to study these effects in a unified and systematic way.

Insights afforded by RUT as a unified framework for comparing sources of preference data extend to many other sources of preference data and types of research applications. For example, there have been a number of comparisons of models estimated from revealed (RP) and stated preference (SP) data sources, such as Ben-Akiva and Morikawa (1990) and Hensher and Bradley (1993), who compared RP and SP transport mode choice data; Louviere, Fox and Moore (1993) who compared several different RP and SP data sources for vacation trips; Swait, Louviere and Williams (1994), who compared RP and SP data sources for freight shipper choices in three cities; Adamowicz, et al. (1994, 1996), who compared RP and SP sources for water-based recreation and moose-hunting destination choices; and Louviere, Hensher and Swait (1998), who compared several RP and SP data sources (again, to name a few). The hypothesis of preference invariance across data sources generally was supported in these comparisons, although some minor discrepancies were found in some studies.

The preceding represent only a few of many new insights into consumer decision making and choice processes now possible from the RUT paradigm. Furthermore, our discussion suggests that many previously published results might bear re-examination. For example, Louviere, Hensher and Swait (1999, Paper 13) reviewed a large number of published studies in marketing, transportation and environmental and resource economics that reported differences in model parameters or decision processes due to differences in product categories, context effects, latent segments, etc. They showed that in many cases the reported empirical differences most likely were due to differences in error variability (ie, sizes of random components), not real differences in model parameters or statistical effects. Hence, failure to recognise and understand the role of the random component may explain many published results. For example, Ainslie and Rossi (1998) recently reported a number of empirical regularities in choice model parameters estimated from different sources of scanner panel data for different product categories, but did not recognise that they were consistent with and could be explained by the error variability mechanism of RUT (they also did not reference the large literature on the role of the scale parameter).

A Theory Of Preference Regularities/Invariance Based On Rut

The preceding discussion provides a conceptual basis to address the general problem of comparing and possibly combining sources of preference data, whether from conjoint analysis or other sources. That is, suppose a sample of consumers respond to a survey that (among other things) asks them to make choices from a designed set of paired-alternative scenarios that describe a product/service. We call this preference elicitation procedure one (PEP1), and it has an associated design matrix X_1. Suppose we also have data from a second sample of consumers for the same product/service consisting of self-reports about which product/service they last purchased (ie, PEP2). Associated with these self-reported choices is design matrix X_2 representing consumer's perceptions of attribute levels of each product/service. In general, X_1 and X_2 have some common attributes (say, X_{c1} and X_{c2}), plus others that are data source-specific (eg, Z_1 and Z_2, respectively). Thus, $X_1=(X_{c1},Z_1)'$ and $X_2=(X_{c2},Z_2)'$.

We now specify the utility function for each data source in terms of common and data source-specific attributes. For pedagogical simplicity, let the utility expressions be strictly additive in all effects, let the common and data-source-specific attributes have separate error terms and let both data sources have different error components to account for typical statistical issues of measurement errors, omitted variables, wrong functional forms, etc. Denote these error components ζ_1, ζ_2, ζ_{c1}, ζ_{c2}, ε_1, and ε_2, respectively, so that we can write the utility expressions as follows:

$$U_1=\theta_1+[V_{c1}(X_{c1},\beta_1)+\zeta_{c1}]+[W_1(Z_1,\gamma_1)+\zeta_1]+\varepsilon_1 \quad (4)$$

$$U_2=\theta_2+[V_{c2}(X_{c2},\beta_2)+\zeta_{c2}]+[W_2(Z_2,\gamma_2)+\zeta_2]+\varepsilon_2 \quad (5)$$

where quantities V_{c1} and V_{c2} are the utility components of the common attributes with associated parameters (β_1, β_2); W_1 and W_2 are the utility components of the data source-specific attributes with associated parameters (γ_1 and γ_2); θ_1, θ_2 are intercepts that measure average preference levels in each data source; and the error components are as previously defined. Rearranging we have:

$$U_1=\theta_1+V_{c1}(X_{c1},\beta_1)+W_1(Z_1,\gamma_1)+(\zeta_{c1}+\zeta_1+\varepsilon_1) \quad (6)$$

$$U_2=\theta_2+V_{c2}(X_{c2},\beta_2)+W_2(Z_2,\gamma_2)+(\zeta_{c2}+\zeta_2+\varepsilon_2) \quad . \quad (7)$$

The dimensionality of each data source is defined by its PEP. That is, U_1 has two rows per choice set because it is a paired choice task, but the number of rows of U_2 may vary from consumer-to-consumer because the number of brands that each reports were in their choice sets when they made their last purchase can vary between consumers.

A key issue suggested by the above discussion is whether responses obtained from different PEPs, contexts, etc., reflect the same underlying preference processes. That is, are the common utilities $V_{ck}(X_{ck},\beta_k)$ the same, despite being estimated from different PEPs, contexts, etc? In order to address this issue, a formal definition of preference regularity or invariance is needed. As discussed by Louviere, Hensher and Swait (1999), two or more PEPs exhibit preference regularity or invariance if the marginal common utilities for any (k,k') pair of data sources are proportional. That is, $\beta_k \propto \beta_{k'}$, and the constant of proportionality should be $\lambda_k/\lambda_{k'}$.

Strictly speaking, this definition holds only if the common attribute parameters are specified as linear-in-the-parameters. More generally, the marginal rates of substitution ($\partial V_{ck}(X_{ck},\beta_k)/\partial X_{ck}$) must be proportional. The linear-in-the-parameters case is more restrictive, but may be more applicable because most utility functions used in choice model applications are linear-in-the-parameters. A simple way to view preference regularity or invariance in RUT choice models is to graph one vector of estimated common utility parameters against a second. If invariance holds, the graphed points should lie on a straight line that intersects the origin. More generally, if there are two or more vectors of common model parameters, one must serve as a (reference) vector graphed on the X-Axis, with the other vectors on the Y-Axis. In this case, the vectors should plot as a fan of straight lines that intersect at the origin of the graph. The slope of each line is the constant of proportionality relative to the reference vector (ie, $\lambda_k/\lambda_{k'}$).

Put another way, if the parameter vectors plot as a fan of straight lines intersecting the origin (ie, preference invariance holds), they will be linear combinations of one another because they are proportional. Thus, a factor analysis of the vectors of model parameters (rows = parameters, columns = models) should yield a single factor (See Krijnen 1997 for a proof). Confirmatory factor analytic procedures could be used to test this hypothesis, but often there are few parameters and power is lost in aggregation (parameters are means).

Each common attribute utility parameter estimates $\partial V_{ck}(X_{ck},\beta_k)/\partial X_{ck}$; hence, graphs and factor analyses are not statistical tests, but instead are diagnostic aids. That is, one must take errors of sampling and estimation into account to properly test a preference invariance hypothesis that retains full statistical power. For example, one can generalise the test proposed by Swait and Louviere (1993) by treating parameter proportionality as a restriction tested by a Full Information Maximum Likelihood (FIML) procedure. In fact, model parameter proportionality is a very strong requirement, and its stringency increases with the number of attributes. For example, in the two data source case one first estimates separate models from each source, then pools both sources to estimate a single common vector of attribute parameters with the restriction that β's in data source 1 are proportional to β's in source 2. The pooled model will have K-1 fewer parameters if there are K total common β parameters. and twice the difference in the sum of the separate model likelihoods minus the pooled model likelihood is distributed as chi-square with K-1 degrees of freedom.

The null hypothesis in this test is that both β vectors are the same up to rescaling by a constant of proportionality. This hypothesis should be rejected if there are differences in utility functions, choice sets (of alternatives) and/or decision (choice) rules used to select alternatives from choice sets, which might be due to differences in contexts, frames, orders, or any of a large number of other possibilities. Thus, it is both surprising and significant that there have been more than a dozen empirical tests of this hypothesis involving different sources of data collected under different conditions in different places at different times, etc., but few serious rejections. Thus, preference invariance or parameter proportionality (or its inverse, error variance proportionality) seems to account for a very wide range of differences in data sources. In fact, the success of this simple mechanism in explaining differences in empirical results in many published and unpublished cases now places the onus on consumer behaviour, judgment and decision making and behavioural decision theory researchers to demonstrate that *it cannot explain their findings*. Similarly, the onus is on conjoint analysis researchers and practitioners to demonstrate that *there is a compelling reason not to estimate and test RUT-based choice models* instead of traditional non-RUT flavours of conjoint given the consistent empirical success of the RUT paradigm.

It also is important to recognise that preference regularity or invariance does not require the alternative-specific constants (ASCs) of choice models from different PEPs to be comparable. That is, apart from the obvious fact that preference invariance is defined in terms of common attributes, ASCs are location parameters of the random utility component, hence not associated with attributes. ASCs capture average effects of omitted variables, which can vary between data sources. For example, in the case of the MNL choice model, including all ASCs in a model guarantees that aggregate predicted marginal choice distributions will exactly equal observed distributions (Ben-Akiva and Lerman 1985). Although other choice models may not have this property, ASCs still will be specific to data sources; hence should not be included in tests of preference regularity.

Generality Of The Luce And Suppes Ranking Theorem

The preceding discussion provides a basis for understanding why and how many sources of preference data can be compared on a level playing field. Thus, we now briefly discuss how some common preference measures can be transformed to be consistent with RUT, such as the following:

1. ***Discrete choices of one option from a set of options.*** Discrete choices are observed in so-called "choice-based conjoint" experiments, more general choice experiments, self-reports of last brand purchased, self-reports of most preferred brand from a list of brands, and many other possibilities.

2. ***"Pick-any" choices that indicate preference orderings***, such as "considered" or "liked" brands, and/or questions where more than one listed item can meet an implied threshold. These include responses to lists of brands or conjoint and/or tasks in which than one option meets some threshold preference value, etc.

3. ***Complete or incomplete preference rankings*** from ranking a list of brands in order of preference, ranking conjoint profiles, ranking options in choice experiments, etc. A ranking is "complete" if all options are ranked, and "incomplete" if only a subset are ranked (eg, "top four," "best and worst," "those that actually would be purchased," etc).

4. ***Preferences expressed on (equal interval) category rating scales*** or other scales assumed to yield cardinal measures (eg, magnitude estimation or ratio production scales). Examples include ubiquitous brand evaluations, responses to traditional conjoint analysis tasks and many more possibilities.

Each of the above PEPs provides data that can be transformed to be consistent with RUT. Comparisons of such data sources are important in their own right. That is, if a particular utility specification and/or choice process is hypothesised to underlie a particular response task, this hypothesis also must hold for any arbitrary monotone transformation of the response data into implied discrete choices, as we now explain. Suppose one posits that a utility function is additive and/or an attribute has a certain marginal rate of substitution (or other germane process hypotheses), and tests this by estimating a model from responses that are assumed to be ordinal or higher in measurement level (eg, interval or ratio measures). To generalise, the hypothesis also must hold for any arbitrary monotone transformation of the response data, and if not satisfied it cannot be generalised and results will be unique to certain measurement types and/or response modes.

For example, suppose one estimates a utility specification and associated marginal rates of substitution (MRS's) from a conjoint ratings task. If the ratings from that task are transformed to be consistent with discrete categorical, RUT-based choice models, one must obtain the same inferences about specification and MRS's from an analysis of the latter data for the former hypothesis to be generalised. If analytical results differ significantly, inferences from analyses requiring stronger metric assumptions (eg, ratings require equal interval, or at least ordinal, as opposed to categorical, discrete choice assumptions) would be rejected in favour of results based on weaker assumptions. We now demonstrate how to transform each data type previously discussed to be consistent with RUT (See Luce and Suppes 1965; Ben-Akiva and Lerman 1985):

1. Subjects in discrete choice tasks indicate one preferred or chosen option from a set of N total options. Choice sets are constructed such that chosen options are coded 1 and the N-1 unchosen options are coded 0. There may be only one choice set per subject (eg, consumers self-report their last purchased brand or most preferred brand in a list of N); or there may be several choice sets (eg, subjects in discrete choice experiments indicate one choice from N options in each of several choice sets).

2. Subjects in "pick-any" tasks indicate that 0, 1 or more of N total options meet a criterion. For example, "consideration" task subjects indicate all brands that they would "consider" from a list of N; or in choice experiments, indicate all options they would "consider" in each choice set of N options (ie, c options are "considered" and N-c are "not"). This implies that the c "considered" options are preferred to the remaining N-c options, which

allows one to construct 0 (if c=0), 1 (if c=1), 2 (if c=2) or more (if c > 1) choice sets such that each of the c options is preferred to the remaining N-c options. In this way one can construct c total choice sets (if c=N no sets can be constructed).

3. Subjects in ranking tasks rank all or some subset, r, of N options. Here the first ranked option is preferred to the remaining N-1 options; the second ranked is preferred to the other N-2 options; and so forth (See also Chapman and Staelin 1982). This "rank-order explosion" procedure can be used to construct up to N-1 choice sets from a set of N rankings, or r-1 choice sets from a subset of r rankings of N options. Such data include lists of N brands ranked in surveys, ranking of N conjoint-type profiles, ranking of N options in experimental choice sets, and many other possibilities.

4. The above procedure (3) can be used to construct choice sets from responses assumed to be cardinal measures, such as rating or magnitude estimation tasks, except that ties may arise if options receive equal numerical responses. If t items are tied, t choice sets should be constructed that include only one of the tied options per set.

In addition to the above transformation(s) of rating and ranking data, RUT models can be derived from ratings or similar preference responses by treating them as ordinal indicators of latent, underlying continua. That is, let the observed response be Y_n (eg, a 1 to 7 category rating response), and let the latent scale be U_{SP}. Then we can write:

$$U_{SP}=\beta_{SP}X_{c,SP}+(\zeta_{c,SP}+\varepsilon_{SP}) \; , \qquad (8)$$

with $\nu_{SP}=(\zeta_{c,SP}+\varepsilon_{SP})$ logistically distributed (location parameter 0 and standard deviation σ_{SP}). The cumulative density function for U_{SP} is (Johnson, Kotz and Balakrishnan, 1995)

$$G_{SP}(u)=\{1+\exp[\lambda_{SP}(\beta_{SP}X_{c,SP}-u)]\}^{-1}, \; -\infty<u<\infty, \text{ and } \lambda_{SP}=\pi\sqrt{3}/\sigma_{SP}. \qquad (9)$$

We have to relate the latent scale to the observed responses (Y_n) to specify a RUT model. This is achieved by noting that if U_{SP} is less than or equal to some value τ_1, the subject answers Y_n=1, which event probability (Eqn. 9) = $G_{SP}(\beta_{SP}X_{c,SP}+\tau_1)$. If U_{SP} lies between τ_1 and τ_2, the probability = $[G_{SP}(\beta_{SP}X_{c,SP}+\tau_2) - G_{SP}(\beta_{SP}X_{c,SP}+\tau_1)]$, and the subject responds Y_n=2, and so forth. Thus, parameters $\tau=(\tau_1, ..., \tau_6)'$ are cutpoints or response category boundaries, such that $\tau_1 \leq \tau_2 \leq ... \leq\tau_6$. In this example, only five cutpoints can be identified, so one (eg, τ_1) must be set to 0. Like all RUT models, the variance (or scale λ_{SP}) of the latent variable U_{SP} cannot be identified and is confounded with parameter vectors β_{SP} and τ (equation 9). The same test procedures discussed above can be used to compare models estimated from such responses with other preference data sources. Thus, conjoint ratings data can be transformed to be consistent with RUT, models estimated from each source of preference data can be compared and rigorous tests of preference invariance or regularity can be performed.

For example, Morikawa (1994) combined and compared RP choice data with preference ratings from a traditional conjoint task. RP data were choices between rail and car for intercity travel in the Netherlands, and SP data were graded paired comparisons for the same context. Morikawa tested parameter invariance between data sources and found that the model parameters were proportional

(95% confidence level). The estimated constant of proportionality (ie, $\lambda_{SP}/\lambda_{RP}$) was 0.27, which implies an SP error variance approximately 4 times larger than the RP variance.

Now we use three case studies to illustrate the generality of RUT and how it can be applied to research problems involving conjoint and choice experiment data.

Empirical Case Studies

Case 1: Complex models from simple conjoint choice experiments

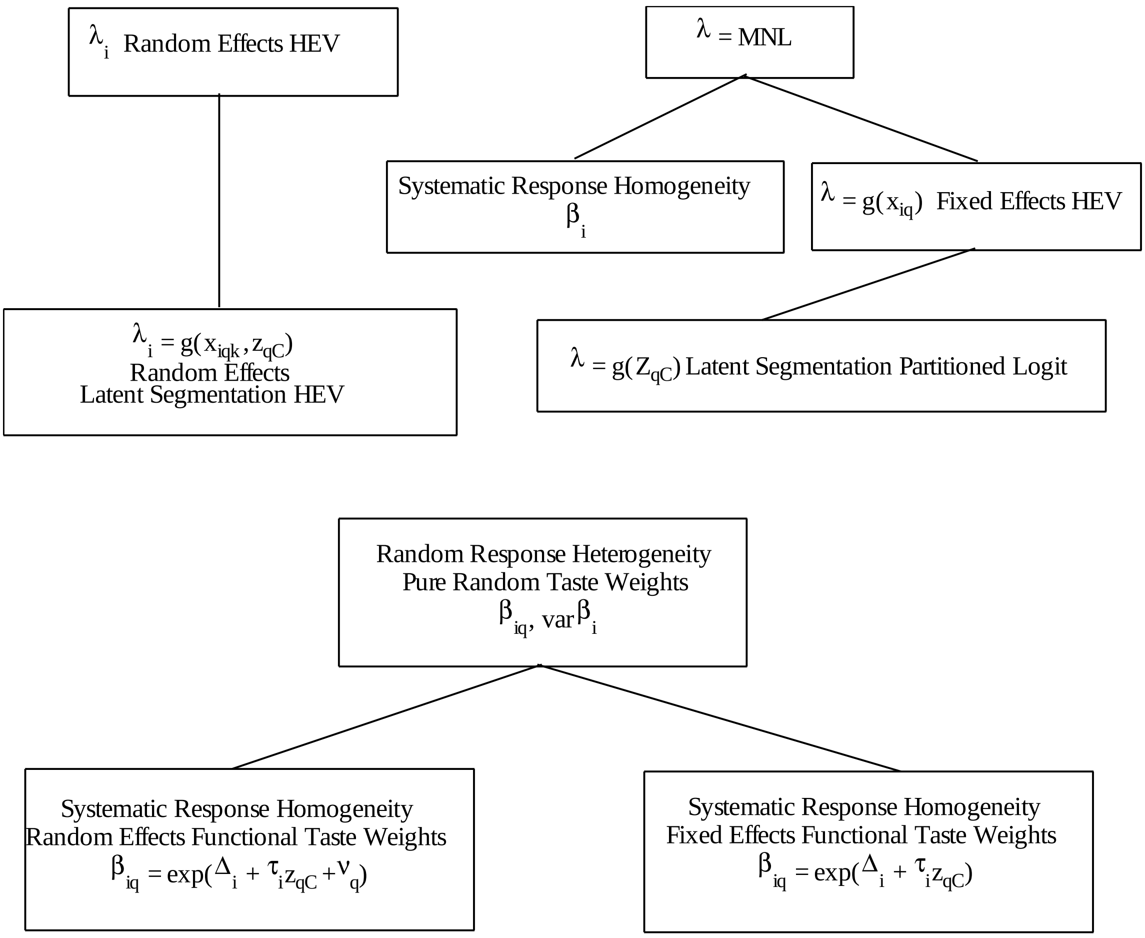
This case study examines differences in marginal rates of substitution derived from simple and complex choice models, the results of which suggest caution in relying on simple choice models like MNL. Indeed, much progress has been made in relaxing the IID error assumptions that underlie simple models like MNL and Identity Probit, although some complex models are little more than mere statistical descriptions devoid of behavioural theory. Moreover, few complex models can forecast future behaviour because they include factors that cannot be forecast easily (if at all) and/or they are merely reduced form approximations of dynamic processes (Erdem and Keane 1996).

Worse yet, few advanced statistical choice specifications can be or have been used to model the full behavioural system of trial and repeat choices, volume decisions conditional on choice and/or inter-purchase time choices, etc. Thus, statistical and mathematical complexity is not a substitute for sound theory and rigorous thinking about process. Indeed, this case illustrates that a rush to complexity may be ill-conceived because recent Monte Carlo work by David Bunch (reported in Meyer, et al., 1999) suggests that numbers of observations needed to satisfy asymptotic theory for complex models (eg, MNP) may be many times greater than simple models like MNL (eg, in some cases required sample sizes may exceed available human populations!).

The more complex the unobserved effects, such as variation and co-variation due to contemporaneous or temporal patterns between alternatives, the more likely it will be that one must simplify complex and often 'deep' parameters associated with covariance matrices to estimate models. Science seeks parsimonious and behaviourally meaningful models rather than complex statistical descriptions, which is why one must understand and appreciate model assumptions. For example, most discrete choice models estimated in conjoint analysis and other paradigms can be described by the following assumptions:

- A single cross-section (no lagged structure);
- Non-separation of attribute utilities from other component effects capturing the role of explanatory variables in utility expressions (due to confounds with scale);
- Constant scale parameters across alternatives (constant variance assumption);
- Random components that are not serially correlated (See Morikawa 1994)
- Fixed utility weights; and
- No unobserved heterogeneity.

A hierarchy of models has evolved that relax some of the above assumptions (Figure 2), and Case 1 focuses on refining the behavioural structure of choice models that treat variance-scale ratio parameters (ie, inverse of random component variances) as real behavioural processes. Specifically, we concentrate on three models: 1) random and fixed effects Heteroskedastic Extreme Value (HEV), 2) Random Parameter or Mixed Logit (RP/ML) and 3) Multi-Period Multinomial Probit (MPMNP).



z_{qC} = A vector of agent-specific characteristics and/or contextual variables

x_{iq} = A vector of alternative-specific attributes which vary across the agents

Multi-period Multinomid Probit
with IID or non-IID
autoregressive errors
unobserved heterogeneity (random effects)
inter-alternative correlation

Figure 2: Taxonomy of Behaviourally Progressive Models

1.1 Heteroskedastic Extreme Value (HEV) Models - Random Effects

If random component variances differ across alternatives, constant variance models will over- or under-estimate the indirect utility effects. HEV allows variance differences but retains zero inter-alternative correlations, hence variance scale ratios (hereafter λ's) can vary across alternatives (probit analogs involve normal distributions). HEV relaxes the IID property of MNL by allowing different variance scale-ratios (λ's) for alternatives, which in turn allows differential substitution among all pairs of alternatives. That is, changes in alternative *l*'s utility affect utility differences in *i* and *l*, but the utility changes are affected by *i*'s λ value. Specifically, the smaller *i*'s λ value, the less effect utility difference changes have, and the smaller the effects on the probability of choosing *i*. Bhat (1995, 1997a), Allenby and Glinter (1995) and Hensher (1997, 1998a) discuss HEV.

1.2 Random Parameter Logit (RPL) or Mixed Logit models

Unlike MNL, RPL treats one or more utility parameters and/or alternative-specific constants as random parameters, the variance(s) and mean (s) of which are estimated. RPL will produce non-IID choice outcomes if random variation of individual parameters induces correlations among the utilities of alternatives (Bhat 1997, McFadden and Train 1996). RPL, or 'mixed logit,' is a mixture of a Gumbel distribution for the error component and a normal distribution for the utility parameters (Train, 1997) and can account for cross-correlation among alternatives. Revelt and Train (1996), Bhat (1996), McFadden and Train (1996) and Brownstone et al (1997) discuss RPL/Mixed logit models. More recently, Bhat (1997) extended RPL/Mixed Logit by including parameterised covariates (Z_{qk}) in the utility function.

1.3 Multi-Period Multinomial Probit

MultiPeriod-MultiNomial Probit (MPMNP) is the most general way to specify the variances and covariances of the random effects; hence HEV and RPL are special cases. MPMNP can relax most random component assumptions: eg, autoregressive structures, correlations of unobserved effects of alternatives and time periods, unobserved heterogeneity, variance differences in random components, etc. Parameter estimation is more complex for MNP models, and requires Simulated Maximum Likelihood (SML) methods that take pseudo-random draws from the underlying error process (Geweke et al 1994, McFadden and Ruud 1994; Boersch-Supan and Hajivassiliou 1990; Stern 1997) or some form of Bayesian estimation (eg, Wedel, et al., 1999). The pros and cons of Bayesian estimation versus SML are not yet clear, but both have many similarities.

We now illustrate the behavioural implications of the three models by using them to analyse the results of a conjoint choice experiment used to forecast demand for a new high-speed rail system in Australia. The experiment was used to create high-speed rail profiles described by fare class (first, full economy, discount economy and off-peak), frequency of service (every 30, 60, 120 minutes) and parking cost ($2 - $20 per day). The choice task focused on each subject's most recent trip, offering a choice of four high-speed rail options or auto if that trip was made again. Subjects evaluated two or four profiles (355 evaluated two, and 81 other subjects did four scenarios). The number of scenarios is not germane to the case, hence is not discussed further.

We estimated the mean values of non-business travel time savings (VTTS) from each model (Table 1). All models except Model 1 (MNL) were estimated using SML (100 replications). Table 1 reveals substantial and significant differences in VTTS, ranging from a low of $4.63/adult person hours for MNL to $8.37/adult person hour for a more complex model. Such large difference in VTTS can have major implications for go/no-go investment decisions in transport or marketing (which might estimate willingness-to-pay instead). Accounting for cross-alternative correlations and relaxing constant variance assumptions (Table 1) significantly reduce the evident downward bias in MNL mean VTTS estimates. In contrast, unobserved heterogeneity (random effects) and serial correlation have less effect on downward bias. This is consistent with McFadden's (1986) point that individual-level utility estimates rarely should be necessary to capture heterogeneity. As well, these results suggest that heterogeneity may result in less model bias than previously thought in conjoint and choice model applications in marketing.

Table 1: Alternative error processes in models and values of travel time savings

Model	Error Processes	RAN	AR1	MNP	VTTS*	LogL
1	iid across periods, iid across alternatives	0	0	0	**4.63**	-1067.9
2	iid across periods, correlated across alternatives	0	0	1	**6.46**	-1050.71
3	random effects, iid across alternatives	1	0	0	**5.22**	-765.01
4	random effects, correlated across alternatives	1	0	1	**6.88**	-759.57
5	AR1 errors, iid across alternatives	0	1	0	**4.98**	-811.46
6	AR1 errors, correlated across alternatives	0	1	1	**7.87**	-770,38
7	random effects + AR1 errors, iid across alt's	1	1	0	**5.40**	-775.68
8	free variance, random effects, iid across alts	1	0	1	**8.37**	-759.71
9	free variance and iid across periods	0	0	1	**7.64**	-1040.25
10	free variance, iid across periods, correlated across alts	0	0	1	**8.06**	-1044.3
11	free variance, random effects, AR1 errors, correlated across alt's	1	1	1	**7.09**	-759.04

* Dollars per adult person hour, - = not able to identify an appropriate model

Case 2: Extension to multiple conditional choices and combining data sources

Many behaviours of interest involve conditional decisions and/or what can be described as "coping strategies" to minimise hassles or losses. For example, commuters can cope with congestion by changing departure times and/or modes; travel mode choice models that ignore departure time choices will overestimate peak hour queuing and delays, increasing the implied value of new road investments. Billions of dollars are at stake in such investments; hence businesses and society have vested interests in development and application of accurate models of such choices. Similarly, investments in new products and services, product enhancements and extensions, etc, may risk many millions of dollars. Thus, firms also should seek more accurate models of choice processes, which often will involve more than the simple choice of a particular brand. Instead, they may involve conditional choices of brands, purchase quantities, inter-purchase timings and the like.

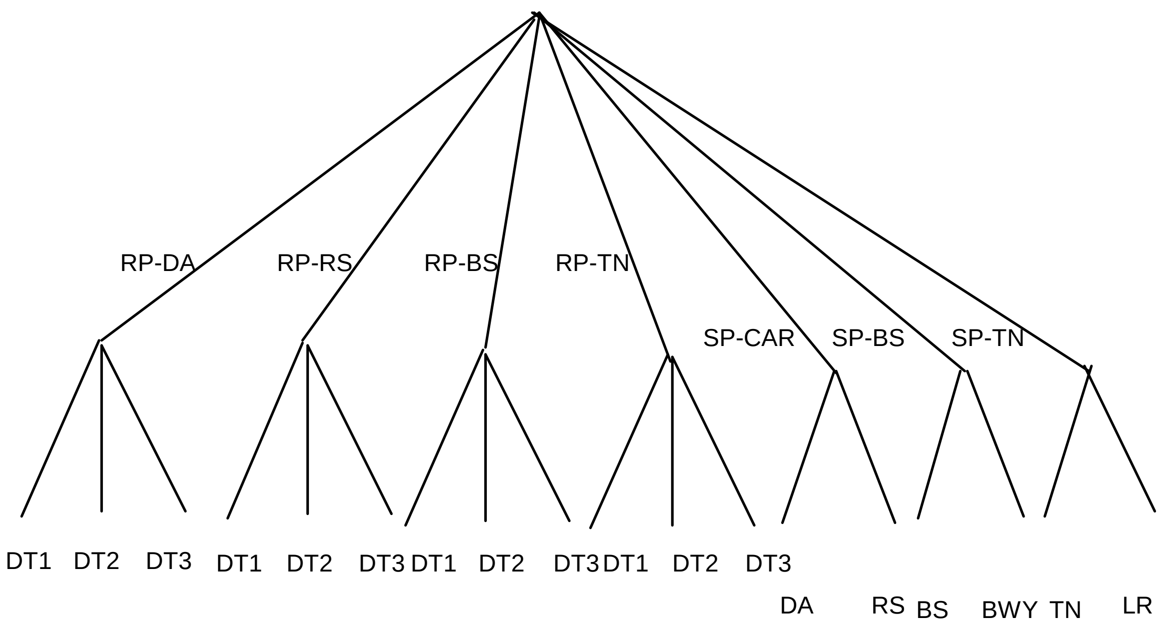
Case 2 illustrates the development and estimation of models of joint choice of departure time and travel mode. The case also illustrates how one can combine a choice (or stated preference, SP) experiment with real choice (or revealed preference, RP) data. The SP task has six travel options: Drive alone, Ride share, Bus, Busway, Train and Light rail (Figure 3). The Bus, Busway, Train and

Light rail options are described by five attributes (total in vehicle time, service frequency, closest stop to home, closest stop to destination, and fare). Drive alone and Ride share options are described by five or six attributes (travel time, fuel cost, parking cost, travel time variability + departure time and toll charge for toll roads).

All attributes were assigned three levels, and a choice experiment was designed by treating all attributes as a collective factorial and selecting an orthogonal fraction from it (Louviere and Woodworth 1983). The selected fraction permits estimation of all main effects within alternatives, plus two-way bilinear interactions for both car options and generic two-way bilinear interactions for public transit options; and produces 81 choice sets, which were blocked into 27 versions of three choice sets due to task complexity. It is worth noting that the nature and number of attributes differ between options, posing design and implementation issues for traditional conjoint (Louviere 1988).

Nested Logit was used to model the departure time and mode choices to allow differential error variances and correlations between subsets of alternatives. RP options included Drive alone, Ride share, Train and Bus (toll and no toll options were combined); and SP options added two 'new' modes (buses on dedicated roads, or “busways”, and light rail), and three departure times based on local data (before 7am, 7.01-8.30 am and after 8.30 am). The tree structure in Figure 3 was one of several investigated based on equivalent variances within partitions, which allows scale (variance) differences between partitions.

Figure 3. The Nested Structure Used in Model Estimation



The model estimation results in Table 2 not only provide new insights into joint choices of departure time and mode, but also allow evaluation of a wide range of policy options to impact these choices. For example, Ride share and Bus unexplained utility components are significantly larger than the other modes that have about equal error variances. This suggests that variability in choice of Ride share and Bus is much greater than for other modes, or the model explains far less in choice. For example, preference heterogeneity is one source of the variability, and these results suggest that it is significantly larger for Ride share and Bus choices.

Table 2: Joint departure time and mode choice model for commuters

Variable	Alts	Units	Estimated Parameter	t-value
Fuel cost (one-way trip)	All	$	.5745	-3.80
Door-to-door time (one-way trip)	rpda,rprs,spda,sprs	mins	-.04224	-2.67
Parking cost	rpda,rprs,spda,sprs	$ per day	-.08978	-3.21
Car availability	rpda,spda		.3789	4.47
Personal income	rpda,spda	$' 000	.00485	1.74
Drive alone constant	rpda		2.0433	1.84
Age <25 years	rprs dt3	years	.3234	1.64
Age 25-34 years	rprs, dt2	years	.1391	1.46
Age 35-54 years	rprs, dt2	years	.0734	1.53
Ride share constant	rprs		1.5412	0.79
Linehaul time	Rpbs,rptn,spbs,sptn, spbwy,splr	mins	-.05409	-5.92
Access and egress time	Rpbs,rptn,spbs,sptn, spbwy,splr	mins	-.03505	-3.47
Managers & Administrators	rpbs,dt1	1,0	8.743	1.83
Clerk	rpbs, dt2 & dt3	1,0	5.785	1.69
Full time employee	rpbs,dt3 & rptn	1,0	.5309	1.57
Age 35-54 years	rptn,dt1	years	-.02754	-1.94
Age 35-54 years	rptn,dt3	years	.01877	1.80
Professionals	rptn,dt2	1,0	-1.724	-3.25
Train specific constant	Rptn		3.088	3.57
Toll cost	spda,sprs	$	-.1466	-2.65
Travel time variability	spda,sprs	mins	-.1269	-2.89
Drive alone constant	Spda		.1289	.29
Ride share constant	Sprs		.1863	.41
Transit frequency	Spbs,sptn,spbwy,splr	mins	-.0267	-3.24
Train specific constant	Sptn		-.3777	-1.22
Light rail specific constant	Splr		-.2633	-.89
Busway specific constant	Spbwy		-.0722	-.67
DT1 specific constant	Rp all dt1's		-.1816	-1.01
DT3 specific constant	Rp all dt3's		.0958	.36
Inclusive Values:				
RP - Drive alone	Rpda		1.075	2.30
RP - Ride Share	Rprs		.278	2.00
RP - Bus	Rpbs		.256	2.01
RP - Train	Rptn		1.00	fixed
SP- Car (Drive alone & ride share)	Spda & Sprs		1.14	3.32
SP – Bus (Bus and Busway)	Spbs & Spbwy		1.07	3.85
SP – Train (Train and Light Rail)	Sptn & Splr		.925	3.91
Log-likelihood at convergence	-2500.82			
Pseudo-r^2	.659			
Sample	2688			

Note: rp=revealed preference, sp=stated preference, da=drive alone (for all departure times), rs=ride share (for all departure times), bs=bus (for all departure times), tn=train (for all departure times), bwy=busway, lr=light rail, dt1=departure time up to 7am, dt2=7.01-8.30am), dt3= after 8.30am. The sample was choice- based sample with RP choice set weights, respectively = .120, .064, .041, .023, .257, .181, .053, .060, .106, .047, .014, .034; all SP weights = 1.0. Estimation by the method of Simulated Maximum Likelihood.

All attributes have the expected signs in both RP and SP data sources, and several individual difference effects also are significant (Table 1), which while common in transport applications of SP methods, is less so in marketing applications (but see Guadagni and Little 1983). The model in Table 3 is a very significant improvement over MNL, suggesting that accounting for variance and individual differences greatly improves MNL and offers other strategic insights. The parameter estimates associated with inclusive values are scale parameters for each partition of the tree. For example, scale parameters for RP-Ride share and RP-Bus (respectively, .278 and .256) imply less random component variance than RP-drive alone and RP-train (respectively, 1.075 and 1.00). We also note that in our experience scale parameters for subsets of RP and SP options often are more similar to one another than scale parameters within RP or SP choice sets.

MNL models cannot account for these types of behavioural differences, suggesting many potential applications for models that can implement the framework in Figure 1: eg, joint brand and quantity choices, joint brand, quantity and inter-purchase time choices, brand and quantity choices conditional on category choice, etc. Thus, Case 2 may seem simple, but it serves to introduce and illustrate more complex tasks and model possibilities.

Case 3: Parameterising the error component in choice experiments

Case 3 illustrates the design and analysis of a fairly complex SP task. The research objective was to develop a model of weekend recreation accommodation choices in and near National Forests in two US states (Missouri and Arkansas) and estimate the likely effects of a wide range of policy changes on campgrounds managed by the US Forest Service. The study was funded by the USDA Forest Service (North Central Forest Experiment Station, Urban Forestry Project, Chicago), who wanted to consider policy changes not previously implemented or investigated systematically. A very large number of actual choice options and resource constraints precluded parallel RP data collection; hence, we discuss only the SP experiment and associated choice models.

Accommodation choices included four National/State Forest options (two campgrounds, rustic cabins and rustic lodges), hotels/motels in nearby towns/villages or staying home and/or doing something else. Each option was described by different attributes, with some attributes common to all options (attributes listed in model estimation results in Tables 3 and 4).

Table 3: Baseline MNL model results for weekend recreation

STATISTICS:

Groups (sets) = 4113
Cases = 24678
Free Parameters = 93
Log LL(B) = -6961.90; Log LL(0) = -7369.51; -2[LL(0)-LL(B)] = 815.221
Rho-Squared = 0.0553; AIC Rho-Squared = 0.0427

	Parameter Estimate	Asymptotic t-Ratio		Parameter Estimate	Asymptotic t-Ratio
Campground	0	—	FishingC	0.0524	1.2
Rustic Cabin	-0.2142	-4.2	BoatingC	0.1723	3.8
Rustic Lodge	-0.6636	-10.7	HikingC	0.1825	4.0
Hotel/Motel	0.0665	1.5	PlayGrdC	0.252	5.5
SEMiss	0.0135	0.3	PicnicC	0.0717	1.6
SWMiss	0.185	3.7	TwoWeeksC	0.0886	2
SCMiss	-0.0145	-0.3	NumCabC	0.024	0.1
Vicinity	-0.0001	0.0	EntryFeeLC	0.0112	0.2
Rugged	-0.0289	-1	CampFeeLC	0.0229	0.5
Lake	0.1633	5.5	SEMissL	-0.2539	-5.5
RegXLake	-0.0354	-1.2	SWMissL	-0.1893	-2.0
VisSep	0.032	1.1	SCMissL	0.1765	2.0
Quiet	0.1564	5.2	VicinityL	0.1502	1.7
VisXQui	0.0377	1.3	RuggedL	0.0103	0.2
TownL	-0.0276	-2.1	LakeL	-0.0271	-0.5
TownQ	-0.041	-1.4	RegXLakeL	0.3169	5.9
Tents	-0.0844	-2.9	TownLL	-0.0053	-0.1
ShoNear	0.215	4.4	TownQL	-0.0201	-0.8
ShoCent	0.174	3.5	BBQL	-0.0925	-1.7
NoSho	-0.1637	-3.1	FishingL	-0.0129	-0.2
ElecHkUp	0.0695	2.3	BoatingL	0.2889	5.4
WaterHkUp	0.0245	0.8	HikingL	0.2352	4.4
BBQ	0.0622	2.1	PlayGrdL	0.2673	5.0
Fishing	0.2258	7.5	PicnicL	0.1005	1.9
Boating	0.1414	4.8	TwoWeeksL	0.0714	1.4
Hiking	0.0614	2.1	NumRoomL	0.0057	0.1
PlayGrd	0.0321	1.1	EntryFeeLL	-0.0272	-0.5
Picnic	0.0179	0.6	CampFeeLL	0.0596	1.1
Friday	-0.0034	-0.1	SEMissH	-0.371	-6.7
Anytime	-0.0839	-2.2	SWMissH	-0.0471	-0.7
TwoCG	0.0184	0.6	SCMissH	0.1713	2.5
Sites75	0.0021	0.1	VicinityH	0.0026	0.0
TwoX75	-0.02	-0.7	RuggedH	0.0716	1.7
NoFee	-0.0076	-0.4	LakeH	-0.0608	-1.5
EntryFeeL	-0.0547	-1.3	RegXLakeH	0.1481	3.6
EntryFeeQ	-0.0142	-0.6	FishingH	0.0158	0.4
CampFeeL	-0.0458	-3.5	BoatingH	0.0953	2.3
SEMissC	-0.0228	-0.8	HikingH	0.108	2.6
SWMissC	0.0417	0.5	PlayGrdH	0.0475	1.2
SCMissC	0.1021	1.4	TwoWeeksH	0.0383	0.9
VicinityC	0.0088	0.1	NumHotelH	-0.0151	-0.4
RuggedC	-0.0242	-0.5	CampFeeLH	-0.055	-1.3
LakeC	0.0514	1.1	CampxTowns	-0.093	-2.3
RegXLakeC	0.2105	4.7	CampxRustic	-0.1098	-3.5
TownLC	0.0781	1.7	CampxWithRoads	-0.015	-0.4
TownQC	-0.0553	-2.7	CampxNoRoads	0.5489	1.6
BBQC	-0.0439	-1.0	CampxWildrness	0.6725	2.7

Table 4: HEV Model Results [Scale=f(Recreation Opportunity Scale)] for weekend recreation

STATISTICS:

Groups (sets) = 4113
Cases = 24678
Free Parameters = 101
Log LL(B) = -6748.85; Log LL(0) = -7369.51; -2[LL(0)-LL(B)] = 1241.32
Rho-Squared = 0.0842; AIC Rho-Squared = 0.0705

Attribute/Model Effect	Parameter Estimate	Asymptotic t-Ratio	Attribute/Model Effect	Parameter Estimate	Asymptotic t-Ratio
ASC-Campground	0.000	—	FishingC	0.025	1.1
ASC-Rustic Cabin	-0.277	-4.3	BoatingC	0.079	3.5
ASC-Rust. Lodge	-0.661	-4.8	HikingC	0.076	3.4
ASC-Hotel/Motel	-0.759	-14.7	Play GroundC	0.137	5.2
SE. Missouri	-0.002	-0.1	Picnic areasC	0.014	0.6
SW. Missouri	0.079	4.4	Reserve 2 WksC	0.037	1.7
SC. Missouri	-0.008	-0.5	No. CabinsC	-0.007	-0.3
NF in vicinity	-0.005	-0.5	Entry FeeLC	0.012	0.6
Rugged	-0.009	-0.9	Camping FeeLC	0.002	0.1
Lake	0.060	5.8	SE. MissouriL	-0.120	-4.8
Rugged x Lake	-0.013	-1.4	SW. MissouriL	-0.075	-1.3
Visually Separate	0.016	1.6	SC. MissouriL	0.053	1.0
Quiet	0.066	6.2	NF in vicinityL	0.095	1.8
VisSep x Quiet	0.016	1.7	RuggedL	0.032	1.0
Distance TownL	-0.006	-1.3	LakeL	-0.019	-0.6
Distance TownQ	-0.013	-1.3	Rugged x LakeL	0.164	4.4
Tents	-0.023	-2.3	Distance TownLL	0.006	0.2
Shower Near	0.092	5.1	Distance TownQL	-0.004	-0.3
Shower Cent	0.060	3.5	BBQsL	-0.057	-1.8
No Shower	-0.070	-3.8	FishingL	0.021	0.7
ElecHkUp	0.019	1.9	BoatingL	0.165	4.3
WaterHkUp	0.012	1.2	HikingL	0.126	3.7
BBQs	0.023	2.2	Play GroundL	0.138	3.9
Fishing	0.085	7.4	Picnic areasL	0.065	1.9
Boating	0.051	4.8	Reserve 2 WksL	0.050	1.6
Hiking	0.032	3.1	No. RoomsL	-0.010	-0.3
Play Ground	0.012	1.1	Entry FeeLL	-0.026	-0.8
Picnic Area	0.001	0.1	Camping FeeLL	0.010	0.3
Arrive Friday	-0.008	-0.5	SE. MissouriH	-0.214	-4.9
Arrive Anytime	0.006	0.5	SW. MissouriH	-0.007	-0.1
2 camp grounds	0.001	0.1	SC. MissouriH	0.099	1.6
75 Sites	-0.001	-0.1	In vicinityH	-0.018	-0.3
2 campg x 75 sites	-0.010	-1.0	RuggedH	0.041	1.1
No Fee	-0.004	-0.7	LakeH	-0.039	-1.0
Entry FeeL	-0.031	-2.2	Rugged x LakeH	0.110	2.9
Entry FeeQ	-0.003	-0.3	FishingH	-0.024	-0.6
Camping FeeL	-0.014	-3.3	BoatingH	0.079	2.1
SE. MissouriC	-0.008	-0.8	HikingH	0.077	2.0
SW. MissouriC	0.033	0.9	Play GroundH	0.052	1.4
SC. MissouriC	0.048	1.3	Reserve 2 WksH	0.027	0.7
NF in vicinityC	-0.017	-0.4	No. HotelsH	-0.042	-1.1
RuggedC	-0.007	-0.3	Camping FeeLH	-0.026	-0.7
LakeC	0.021	1.0	Camping x Towns	-0.095	-2.5
Rugged x LakeC	0.081	3.6	Camping x Rustic	-0.143	-4.2
Distance TownLC	0.035	1.6	Camping x Roads	0.198	5.8
Distance TownQC	-0.018	-1.8	Camp' g x 0Roads	0.091	2.8
BBQsC	-0.016	-0.7	Camp' g x Wilds	0.081	3.3

SCALE FUNCTION PARAMETERS

Main effects (variables) in error variance function	Parameter Estimate	Asymptotic t-Ratio
Campground	0.0607	3.4
Rustic Cabin	1.0997	18.1
Rustic Lodge	0.5296	6.0
Hotel/Motel	0.4168	3.7
Prefer Towns	0.0000	—
Prefer Rustic	-0.1062	-4.7
Prefer WithRoads	0.0834	3.8
Prefer NoRoads	-0.0042	-0.2
Prefer Wilderness	0.0640	2.6

The SP task was designed by treating all attributes of all accommodation choice options (two campgrounds, rustic cabins, rustic lodges and hotel/motel/bed & breakfasts) as a common factorial ((2^{11} x 4^8) x (2^{11} x 4^8) x (2^{13} x 4^3) x (2^{13} x 4^3) x (2^9 x 4) = 2^{57} x 4^{23} (x 8)), from which we selected an orthogonal fraction to make 128 choice sets (Louviere and Woodworth, 1983; Louviere, Hensher and Swait, 1999). Choice sets were blocked into 16 versions of 8 choice sets using the additional 8-level factor. A university survey research center pre-recruited a random sample of Missouri residents by phone, and mailed the survey to those agreeing to participate (no incentive provided), randomly assigning each to one version of the experiment. This design approach was used to permit estimation and comparison of a variety of complex model forms (eg, Nested Logit, HEV, Mother Logit, MNP, etc.).

The survey contained questions additional to the choice experiment (eg, recent outdoor recreation behaviour, types of activities preferred, visits to National Forests and selected socio-demographics) not discussed for space reasons. The analysis and model comparison described below is based on approximately 520 subjects (some subjects omitted 1-2 choice sets; hence the approximation). Data were weighted to equalise sample sizes in the 16 versions (eg, let sample size average 32, with version 1 =20 and version 2 = 45 subjects; then versions 1 and 2 data are weighted by 1.6 and 0.71, respectively), which avoids over- or under-emphasis of design matrices associated with particular versions that can lead to biased parameters.

Again, due to space, we discuss only a baseline MNL model and a parameterised HEV model that allows different random component variances for individuals and choice alternatives. Individual variance differences are a function of preferences for types of recreation environments (from primitive wilderness to developed, modern urban areas; an environment is coded 1 if subjects consider it and -1 otherwise). This variance component model is justified by the fact that many marketing activities and/or policies can impact not only mean utilities or preferences but also their variances (Swait and Louviere 1993; Louviere, Hensher and Swait 1999; Meyer, et al. 1999). Thus, the HEV model in Table 5 allows random component variance differences in both choice options and individuals.

The estimation results clearly favour the HEV model over MNL, even though MNL allowed individual differences in choices to be a function of recreation environment preferences (213 LL point difference for relatively few additional effects). Both models contain alternative-specific attribute effects: the first set of effects without capital letters behind them are generic effects for the two campgrounds, additional effects have capital letters ("C" for cabins, "L" for lodges and "H" for hotels/motels/B&Bs). Recreation environments for which respondents expressed preferences are:

1. Developed urban and modern (cities and towns)
2. Rural, small towns, resorts, some development
3. Rustic, rural, limited development like cabins, lodges, farms, ranches
4. Natural, roads, trails, limited development like isolated cabins & homes
5. Natural, no roads, no established development & marked foot & horseback trails
6. Wilderness, undeveloped, foot and horseback access only, primitive trails

In the interests of brevity, we concentrate on the HEV scaling function results, and ignore most attribute results, except to note that each accommodation option has unique (ie, alternative-specific) attribute effects. For example, subjects who chose campgrounds preferred quiet SW Missouri locations on lakes with fishing and boating; but those who chose cabins, lodges and hotels/motels/B&Bs prefered rugged NW Arkansas locations on lakes with playgrounds, boating and hiking.

The scale (inversely proportional to variance) results reveal that rustic cabins have the least average error variance, "stay at home" and campground have the most average variance, and rustic lodges and hotels/motels/B&Bs have intermediate variances. Those preferring natural environments with roads (#4 above) or undeveloped wilderness (#6 above) exhibit more choice variability than those preferring rustic, rural areas with limited development (# 3 above), followed by those preferring developed modern areas (#1) or natural areas with no roads (#5). These results suggest that it will be harder for policies aimed at managing rustic cabins to produce consistent results for this sample, but campground management policies should yield more consistent results, particularly for those who prefer the rustic, rural environments that characterise Missouri National Forests. For example, those who chose campgrounds tended to prefer SW Missouri National Forest areas on lakes with visually separated and quiet camp sites, fishing, boating and hiking, RV sites, electrical hookups, conveniently located shower facilities and barbecues. They also seem fairly sensitive to both entry and camping fees, which suggests that the costs of providing such facilities should be considered carefully against likely revenues before deciding which strategy to pursue in each National Forest Division.

Conclusions

This paper sought to expand the domain of conjoint analysis techniques by placing them within the more general framework of random utility theory (RUT) based stated preference methods. Thus, we explained that RUT provides a general theoretical framework to design and analyse simple and complex SP experiments to capture a wide array of interrelated behavioural phenomena of interest to marketers, transport planners, environmental economists and many other fields. Furthermore, we noted that RUT also provides a key theoretical link to real behaviour that allows one to pool SP and RP data sources, or more generally any data sources consistent with RUT. Furthermore, RUT provides a level playing field by which models and model results can be compared and rigorously tested.

We illustrated these ideas in three case studies that show that even simple RUT-based SP experiments can yield quite complex models; and complex SP experiments can provide new and different insights into the behaviour of the random component of utility. In fact, the latter case suggests that researchers should consider if the factors they vary in experiments and/or context or experimental differences impact only mean utilities (the traditional object of research interest), but instead also impact the variance in utilities and preferences, a heretofore neglected area of inquiry.

That is, individuals' mean preferences not only differ, but the variability with which they make choices or express preferences also can differ; and differences in variability can matter strategically and substantively (ie, empirically). For example, the more variable an organisation's implementation of a policy (eg, service variability), the more uncertainty individuals have about its true mean, and hence the more variable their response may be to the policy itself. Similarly, the harder for consumers to determine a policy outcome, the more uncertain they will be about what to expect; hence the more likely they will be to stay with what they know. Thus, new phone services, bank accounts, etc., that confuse consumers about fees are likely to struggle against competitors who make it easier for consumers to evaluate fees or established, "safe" competitors.

Likewise, some consumers are inherently (for reasons not yet understood) more variable in their choices and preferences than others. The more consistent a consumer's preferences, the more likely that a policy targeted at her will yield the effects suggested by research. Thus, opportunities exist to develop fundamental knowledge about variability in preferences and choices by developing theory and/or establishing empirical regularities that allow us to understand what drives differences in variability within and between choice options and individuals.

As McFadden (1986) noted, there are few instances in which one needs individual-level utility results to develop useful models of choice behaviour. More complex and behaviourally meaningful RUT-based choice models now provide powerful insights and ways to model and predict choice behaviour. RUT provides a rich and comprehensive theoretical framework with which to develop models, combine sources of preference and choice data, obtain behavioural insights and make more accurate forecasts. This paper introduced these topics and used case examples to illustrate their potential to broadening our knowledge and ability to model and forecast complex behavioural systems.

References

Adamowicz, W.L., Louviere, J. and Williams, M. (1994) “Combining Stated and Revealed Preference Methods for Valuing Environmental Amenities” *Journal of Environmental Economics and Management*, 26:271-292.

Adamowicz, W.L., Swait, J., Boxall, P., Louviere, J. & Williams, M. (1996) ‘Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation’, *Journal of Environmental Economics and Management*, 32:65-84.

Allenby, G. and Ginter, J. (1995) The effects of in-store displays and feature advertising on consideration sets, *International Journal of Research in Marketing*, 12, 67-80.

Ainslie, A. and Rossi, P.E. (1998) “Similarities in Choice Behavior Across Product Categories,” *Marketing Science*, 17 (2), 91-106.

Ben-Akiva, M E and Lerman, S (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand* Cambridge: MIT Press.

Ben-Akiva, M. and Morikawa, T. (1990) “Estimation of Switching Models From Revealed Preferences and Stated Intentions” *Transportation Research A*, 24A(6):485-495.

Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H. and Rao, V. (1994) "Combining Revealed and Stated Preferences Data" *Marketing Letters: Special Issue on the Duke Invitational Conference on Consumer Decision-Making and Choice Behavior*, 5(4).

Ben-Akiva, M., Morikawa, T. and Shiroishi, F. (1991) “Analysis of the Reliability of Preference Ranking Data” *Journal of Business Research*, 23:253-268.

Bhat, C. (1995) A heteroscedastic extreme value model of intercity travel mode choice, *Transportation Research*, 29B (6), 471-483.

Bhat, C. (1996) Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling, forthcoming, *Transportation Research*

Bhat, C. (1997a) Incorporating Observed and Unobserved Heterogeneity in Urban Mode Choice Models, under review, *Transportation Science*

Bhat, C. (1997) Recent methodological advances relevant to activity and travel behavior analysis, Conference Pre-prints, IATBR’97, *The 8th Meeting of the International Association of Travel Behaviour Research*, Austin, Texas, September.

Boersch-Supan, A. and Hajvassiliou, V. (1990) Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models, *Journal of Econometrics*, 58 (3), 347-368.

Bradley, M. and Daly, A. (1994) Use of the logit scaling approach to test rank-order and fatigue effects in stated preference data, *Transportation*, 21 (2), 167-184.

Brownstone, D., Bunch, D. and Train, K. (1997) Joint mixed logit models of stated and revealed preferences for alternative-fuelled vehicles, Conference Pre-prints, IATBR’97, *The 8th Meeting of the International Association of Travel Behaviour Research*, Austin, Texas, September.

Carson, R.T., Louviere, J.J., Anderson, D.A., Arabie, P., Bunch, D.S., Hensher, D.A., Johnson, R.M., Kuhfeld, W.F., Steinberg, D., Swait, J.D., Timmermans, H. and Wiley, J.B. (1994) "Experimental Analysis of Choice, " *Marketing Letters: Special Issue on the Duke Invitational Conference on Consumer Decision-Making and Choice Behavior*, 5(4):351-368.

Chapman, R. and Staelin, R. (1982) “Exploiting Rank Ordered Choice Set Data Within the Stochastic Utility Model” *Journal of Marketing Research*, 19:288-301.

Chrzan, K. (1994) “Three Kinds of Order Effects in Choice-Based Conjoint Analysis, ” *Marketing Letters*, 5 (2):165-172

Elrod, T., Louviere, J. and Davey, K. (1993) "A Comparison of Ratings-Based and Choice-Based Conjoint Models" *Journal of Marketing Research*, 24(3):368-377.

Erdem T. and M.P. Keane (1996) "Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets", *Marketing Science*, 15 (1), 1-20.

Guadagni P.M. and J.D. Little (1983) "A Logit Model of Brand Choice Calibrated on Scanner Data", *Marketing Science*, 2 (3), 203-238.

Geweke, J., Keane, M. and Runkle, D. (1994) Alternative computational approaches to inference in the multinomial probit model, *Review of Economics and Statistics*, LXXVI,(4), 609-632.

Hensher, D. and Bradley, M. (1993) "Using Stated Response Choice Data to Enrich Revealed Preference Discrete Choice Models" *Marketing Letters*, 4(2):139-151.

Hensher, D.A. (1997) A practical approach to identifying the market for high speed rail: a case study in the Sydney-Canberra corridor, *Transportation Research, 31A(6), 431-446*.

Hensher, D.A. (1998a) Establishing a Fare Elasticity Regime for Urban Passenger Transport: Non-Concession Commuters *Journal of Transport Economics and Policy,*32(2), 221-246.

Hensher, D.A., Boersch-Supan, A. and Brewer, A. (1999, in progress*) Alternative behavioural perspectives on interactive agency choice*, Institute of Transport Studies University of Sydney and Department of Economics, Manheim University.

Johnson, N., Kotz, S. and Balakrishnan, N. (1995) *Continuous Univariate Distributions: Volume 2*, 2nd Edition, New York:Wiley.

Johnson, R. (1989) "Making Decisions with Incomplete Information: The First Complete Test of the Inference Model" *Advances in Consumer Research*, 16:522-528.

Keane, M.P. (1997) "Current issues in discrete choice modelling, *Marketing Letters*" 8(3), 307-322.

Krijnen, W.P. (1997) "Some Remarks To Using Single-Factor Analysis As A Measurement Model," Unpublished working paper, Department of Marketing, University of Groningen, June 3.

Louviere, J.J. (1988). *Analyzing Decision Making: Metric Conjoint Analysis*. Sage University Papers Series in Quintitative Applications in the Social Sciences, No. 67, Newbury Park, CA.

Louviere, J. (1994) "Conjoint Analysis" in *Advances in Marketing Research*, R. Bagozzi (editor), Blackwell Publishers.

Louviere, J. and Woodworth, G. (1983) "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data" *Journal of Marketing Research*, 20:350-367.

Louviere, J., Fox, M. and Moore, W. (1993) "Cross-Task Validity Comparisons of Stated Preference Choice Models" *Marketing Letters*, 4(3):205-213.

Louviere, J.J., Hensher, D.A. and Swait, J. (1998) "Combining Sourcs of Preference Data," *Journal of Econometrics*, Special Issue on Marketing and Econometrics, forthcoming.

Louviere, J.J., Hensher, D.A. and Swait, J. (1999) *Stated Choice Methods: Analysis and Applications in Marketing, Transportation and Environmental Valuation,*Cambridge University: Cambridge Press.

Luce, R.D. and P. Suppes (1965) "Preference, utility and subjective probability", in Luce R. D., Bush R. R., and Galanter E. (eds.), *Handbook of Mathematical Psychology*, Vol. III, Wiley (New York).

McFadden, D. (1974) "Conditional logit analysis of qualitative choice behaviour", in Zarembka P. (ed.), *Frontiers in Econometrics*, New York: Academic Press, 105-142.

McFadden, D. (1986) "The Choice Theory Approach to Marketing Research," *Marketing Science*, 5 (4), 275-297.

McFadden, D. (1981) Econometric models of probabilistic choice, in Manski, C.F. and McFadden, D. (eds.) *Structural*

Analysis of Discrete Data with Econometric Applications, Cambridge, MA: MIT Press, 198-272.

McFadden, D. and Ruud, P.A. (1994) Estimation by simulation, *Review of Economics and Statistics*, LXXVI,(4), 591-608.

McFadden, D. and Train, K. (1996) Mixed MNL models for discrete response, *Department of Economics,* University of California at Berkeley.

Meyer, R. (1977) "An Experimental Analysis of Student Apartment Selection Decisions Under Uncertainty," Special Issue on Human Judgment and Spatial Behavior, *Great Plains-Rocky Mountains Geographical Journal*, 6:30-38.

Meyer, R. and Eagle, T. (1982) "Context Induced Parameter Instability in a Disaggregate-Stochastic Model of Store Choice," *Journal of Marketing Research*, 19:62-71.

Meyer, R., Louviere, J., Bunch, D., Carson, R., Delleart, B., Hanemann, M., Hensher, D. Irwin, J. (1999) "Combining Sources of Preference Data" *Marketing Letters,* Special Issue on the HEC Invitational Conference on Consumer Decision-Making and Choice Behaviour, forthcoming.

Morikawa, T (1989) *Incorporating Stated Preference Data in Travel Demand Analysis*, PhD Dissertation, Department of Civil Engineering, M.I.T.

Morikawa, T. (1994) Correcting state dependence and serial correlation in the RP/SP combined estimation method, *Transportation*, 21 (2), 153-166.

Oliphant, K., Eagle, T., Louviere, J. and Anderson, D. (1992) "Cross-Task Comparison of Ratings-Based and Choice-Based Conjoint" presented at the 1992 Advanced Research Techniques Forum of the American Marketing Association, June, Beaver Creek, CO.

Olsen, G.D. and Swait, J. (1995) "Nothing is Important," *Unpublished Working Paper,* Faculty of Management, University of Calgary, Alberta, Canada.

Olsen, G.D., Swait, J., Johnson, R., Louviere, J. (1995) "Response Mode Influences on Attribute Weight and Predictive Ability When Linear Models are Not Certain to be Robust" *Unpublished Working Paper,* Faculty of Business, University of Calgary, Alberta, Canada.

Revelt, D. and Train, K. (1996) Incentives for appliance efficiency: random parameters logit models for households' choices, *Department of Economics*, University of California, Berkeley.

Rushton, G., 1969, Analysis of Spatial Behaviour by Revealed Space Preference, Annals of the Association of American Geographers, 59:391-400.

Simon H.A. (1983). *Reason in Human Affairs.* Stanford University Press, Stanford.

Stern, S. (1997) Simulation-based estimation, *Journal of Economic Literature*, XXXV. December, 2006-2039.

Swait, J. and Louviere, J. (1993) "The Role of the Scale Parameter in The Estimation and Use of Multinomial Logit Models" *Journal of Marketing Research*, 30:305-314.

Swait, J., Louviere, J. and Williams, M. (1994) A sequential approach to exploiting the combined strengths of SP and RP data: application to freight shipper choice, *Transportation*, 21(2):135-152.

Swait, J., Adamowicz, W. and Louviere, J. (1998). Attribute-based stated choice methods for resource compensation: An application to oil spill damage assessment. *Paper presented at the NOAA Workshop on Application of Stated Preference Methods to Resource Compensation*, Washington, DC.

Thurstone, L.L. (1927) "A Law of Comparative Judgment," *Psychological Review,* 34, 273-286.

Train, K. (1997) Mixed logit models for recreation demand, in Kling, C. and Herriges, J. (eds.) *Valuing the Environment Using Recreation Demand Models*, Elgar Press, New York.

Wedel, W., Kamakura, W., Arora, N., Bemmaor, A. Chiang, J. Elrod, T., Johnson, R., Lenk, P., Neslin, S., and Poulsen, C.S. (1999) "Heterogeneity and Bayesian Methods in Choice Modeling" *Marketing Letters,* Special Issue on the HEC Invitational Conference on Consumer Decision-Making and Choice Behaviour, forthcoming.