THURSTONIAN-BASED ANALYSES: PAST, PRESENT, AND FUTURE UTILITIES

ULF BÖCKENHOLT

MCGILL UNIVERSITY

Current psychometric models of choice behavior are strongly influenced by Thurstone's (1927, 1931) experimental and statistical work on measuring and scaling preferences. Aided by advances in computational techniques, choice models can now accommodate a wide range of different data types and sources of preference variability among respondents induced by such diverse factors as person-specific choice sets or different functional forms for the underlying utility representations. At the same time, these models are increasingly challenged by behavioral work demonstrating the prevalence of choice behavior that is not consistent with the underlying assumptions of these models. I discuss new modeling avenues that can account for such seemingly inconsistent choice behavior and conclude by emphasizing the interdisciplinary frontiers in the study of choice behavior and the resulting challenges for psychometricians.

Key words: consistency conditions, identifiability, random utility models, social dependencies.

Introduction

Most psychometric models for the analysis of multivariate discrete choice data are based on the assumption that choices among options are determined by their underlying utilities for the decision-makers. Because of significant computational advances in estimating the parameters of the latent utility distribution, the analysis of multivariate choice data is now well on its way of becoming a routine matter (Böckenholt & Tsai, 2006; Caffo & Griswold, 2006). Even the joint modeling of discrete and continuous choice outcomes is starting to be commonplace (Hanemann, 1984; Gueorguieva & Agresti, 2001). It is therefore appropriate to revisit this methodology and to assess its past, current, and future value in modeling choice data.

Because this paper reflects my personal journey and views on the analysis of choice data, many important subjects as well as contributors are left out, for which I apologize. In particular, this paper will not get into the extensive literature on estimation and inference—careful discussions of these topics can be found in the books by Hensher, Rose, and Greene (2005), Train (2003), and Louviere, Hensher, and Swait (2000). The papers presented at the 2004 Invitational Choice Symposium (Chakravarti, Sinha, & Kim, 2005) provide valuable overviews of many interesting topics in modeling choice behavior. In addition, the addresses of two nobel prize laureates in economics (Kahneman, 2003; McFadden, 2001) contain excellent reviews of this field.

My journey is linked to two papers by Thurstone (1927, 1931) that started the era of psychometric choice modeling. The next section will briefly review the contributions of these papers, which is followed by a discussion of current versions of Thurstonian-based choice models. Subsequently, I will point to new research avenues in extending these models to account for seemingly inconsistent choice behavior. The paper concludes by pointing to the interdisciplinary frontiers of studying choice behavior and the resulting future challenges for psychometricians.

Utility Theories and Choice Data

Most psychometric choice models are based on the notion of decision utility. This approach infers utility from observed choices and in turn uses utility to explain choices. For example, if a person chooses a and rejects b, this behavior is interpreted as indicating that a has a higher utility than b to this person. A number of consistency conditions have been proposed in the literature to characterize choices based on decision utilities. Transitivity is a well-known consistency condition stating that a preference for a over b and b over c implies a preference for a over c.

In psychometrics, Thurstone's (1927, 1931) work proved to be highly influential in promoting an extended version of the decision-utility concept that allowed for stochastic variability in choices. Invoking Fechner's (1860) psychophysical concept of a sensory continuum, Thurstone (1927) argued in his "Law of Comparative Judgment" that choice options can be represented along a utility continuum by random variables that describe the options' effects on a person's cognitive apparatus. Thurstone used randomness as a device to represent factors that determine the formation of preferences but are unknown to the observer. This specification was driven by pragmatic considerations since Thurstone was well aware that he could not distinguish between the positions of whether a choice process is inherently random or determined by a multitude of different factors that may not be measurable. By not specifying the "discriminal" process by which a person "identifies, distinguishes, discriminates, or reacts to stimuli" (1927, p. 369) and defining it to be exogenous to all the factors that can affect a choice process, he broadened tremendously the applicability of this approach to the modeling of choice data. Importantly, he also utilized these ideas for modeling individual differences in choice behavior by arguing that heterogeneity in individual assessments of the same option can be described as realizations of the normal distribution. Introducing Thurstone's work to economics, Marschak (1960) coined the term Random Utility Maximization (RUM) when referring to choice probabilities that result from the maximization of utilities with random elements.

Not unlike the notion of personality traits, a controversial corollary of decision utility is that the measured utilities are predetermined, decomposable, and stable over both time and across situations. Consistency conditions defined for stochastic preferences include stochastic transitivity and contraction and expansion consistency (Block & Marschak, 1960; Falmagne, 1985). Under contraction consistency, if a choice set T is narrowed to U and the options chosen from T are still in U, then no unchosen options should be chosen and no previously chosen options should be unchosen. Under expansion consistency, if a choice set T is expanded to U, then the probability of choosing an option from U should not exceed the probability of choosing the same option from T.

In addition to providing a probabilistic version of decision utility, Thurstone (1931) also promoted the benefits of experimental methods in collecting choice data. By asking respondents to state directly their preferences for a single or multiple sets of options in controlled experimental settings, he showed that it is possible to obtain information about a person's preferences that may not be available by observing choices in the person's daily context. These two data types became known later as "stated preference" and "revealed preference" data.

Thurstone (1931) showed that stated preference data can be used to test for indifference curves which at that time played a pivotal role in consumer demand theory. Initially, economists conceptualized the unit of utility as the "just perceivable increment of pleasure" (Edgeworth, 1881, p. 99) but moved later towards the concept of decision utility and, in this context, adopted Pareto's (1900) work on indifference curves that depict different bundle combinations a person is indifferent to choose among. Interestingly, Thurstone's (1931) experimental investigation of indifference curves was met initially with severe criticism by economists (Georgescu-Roegen, 1936; Wallis & Friedman, 1942) and had little impact on subsequent developments of demand theory. Reasons for the negative reactions included the hypothetical nature of the choice situation,

the apparent empirical difficulties in detecting indifference, and the experimental lack of control for the effect of income and prices. Looking back, however, Thurstone's (1931) study was rather significant because it is the first reported experiment in behavioral economics, and also prompted much future work on conjoint measurement and related methods (Luce & Tukey, 1964).

Thurstone's contributions to experimental paradigms for collecting choice data and to the psychometric foundation of choice models with latent variables, in combination with parallel developments in biometrics, sociometrics, and econometrics, led to a rich set of choice models (Ashford & Sowdon, 1970; McFadden, 1984; Hausman & Wise, 1978). The next section presents a short discussion of these Thurstonian-based models for stated preference data and also points to related developments.

Thurstonian Random Utility Models

Typically, stated choice data are collected in the form of incomplete and/or partial rankings. Incomplete ranking data are obtained when a decision-maker considers only a subset of the choice options. For example, in the method of paired comparison, two choice options are presented at a time, and the decision-maker is asked to select the more preferred one (David, 1988). In contrast, in a partial ranking task, a decision-maker is confronted with all choice options and asked to provide a ranking for a subset of the options. For instance, in the best–worst method, a decision-maker is instructed to select the best and worst options out of the offered set of choice options. Both partial and incomplete approaches can be combined by offering multiple distinct subsets of the choice options and obtaining partial or complete rankings for each of them. Pick any constant-sum, and ordinal versions of paired comparison and rankings are alternative methods for collecting stated preference data (Böckenholt, 1992, 2001a, 2001b; Böckenholt & Dillon, 1997a; Yao & Böckenholt, 1999).

The Response Model

Consider a set of J choice alternatives ($j = 1, \ldots, J$) and n decision-makers ($i = 1, \ldots, n$). Under Thurstone's (1927) random utility approach, a choice of an option j is determined by an unobserved utility assessment, υ_j, that can be decomposed into a systematic and a random part:

$$\upsilon_j = \boldsymbol{\mu}_j + \epsilon_j, \tag{1}$$

where the option mean $\boldsymbol{\mu}_j$ is assumed to stay the same in repeated evaluations of the option, but the random contribution ϵ_j varies from evaluation to evaluation according to some distribution. The stochastic component represents unobserved attributes affecting choice, interindividual differences in the utility assessments, measurement errors, and functional misspecification (Manski, 1977). Clearly, this decomposition is fragile and not easy to perform in empirical applications, especially involving revealed choice data. Initially, (1) was applied to the analysis of paired comparison data obtained from a group of judges who were asked to choose their preferred option for all possible pairs of stimuli. Bock (1958) noted that in this case a two-level representation is needed and extended (1) to include person-specific effects:

$$\upsilon_{ij} = \boldsymbol{\mu}_{ij} + \epsilon_{ij}, \tag{2}$$

where $\boldsymbol{\mu}_{ij}$ represents person's i mean evaluation of option j. Takane (1987) further extended this work by presenting a general class of covariance structures to represent individual differences in the assessments of the choice options.

In general, the choices made by person i for a single choice set of J options can be summarized by an ordering vector $\boldsymbol{r}_i$. For instance, $\boldsymbol{r}_i = (h, j, \ldots, l, k)$ indicates that choice

option h is judged superior to option j which in turn is judged superior to the remaining options, with the least preferred option being k. The probability of observing this ordering vector for person i can be written as

$$\Pr(\boldsymbol{r}_i = (h, j, \ldots, l, k)) = \Pr[(v_{ih} - v_{ij} > 0) \cap \cdots \cap (v_{il} - v_{ik} > 0)]. \tag{3}$$

More compactly, a $(J - 1) \times J$ contrast matrix, $\boldsymbol{C}_i$, can be defined to indicate the sign of the differences among the ranked options. For example, for the ordering vectors $\boldsymbol{r}_i = (j, l, k)$ of $J = 3$, the corresponding contrast matrix takes on the form

$$\boldsymbol{C}_i = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \end{bmatrix},$$

where the three columns of the contrast matrix correspond to the options j, k, and l, respectively.

The distributional assumptions on the ϵ_{ij}'s determine the probability for each response pattern. For example, a multivariate probit model is obtained when assuming that latent judgments $\boldsymbol{v}_i = (v_{i1}, \ldots, v_{iJ})$ of the J options for judge i are multivariate normal with mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}$. Ansari and Iyengar (in press) proposed a more general random-effects representation by specifying the ϵ_{ij}'s semiparametrically using Dirichlet process priors. In the following, the popular normality representation is adopted for ease of presentation. In this case, the probability of observing the ordering vector $\boldsymbol{r}_i$ is obtained by evaluating a $(J - 1)$-variate normal distribution,

$$\Pr(\boldsymbol{r}_i | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}) = \frac{|\boldsymbol{\Gamma}_i|^{-1/2}}{(2\pi)^{(J-1)/2}} \int_0^\infty \cdots \int_0^\infty \exp\left\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\delta}_i)'\boldsymbol{\Gamma}_i^{-1}(\boldsymbol{x} - \boldsymbol{\delta}_i)\right\} d\boldsymbol{x}, \tag{4}$$

where $\boldsymbol{\delta}_i = \boldsymbol{C}_i \boldsymbol{\mu}_i$ and $\boldsymbol{\Gamma}_i = \boldsymbol{C}_i \boldsymbol{\Sigma} \boldsymbol{C}_i'$.

Typically, choice options are presented to judges in multiple blocks, especially when the set of choice options is large. This approach simplifies the judgmental task since only a few options need to be considered at a time. It also facilitates studying whether judges are consistent in their evaluations of the choice options. Consider, for example, a paired comparison task in which a person is asked to choose the preferred option for a subset or all of the pairwise choice sets. Extending (2), the utilities of judge i that underlie the pairwise comparison of options j and k, respectively, can be written as

$$\begin{aligned} v_{ij(k)} &= \boldsymbol{\mu}_{ij} + \epsilon_{ij(k)}, \\ v_{ik(j)} &= \boldsymbol{\mu}_{ik} + \epsilon_{ik(j)}. \end{aligned} \tag{5}$$

The terms $\epsilon_{j(k)}$ and $\epsilon_{k(j)}$ capture random variations in the judgmental process, where $\epsilon_{j(k)}$ denotes the random variation of option j when compared to option k. Thus, according to (3), the choice between j and k is determined by the sign of the latent judgment y_{ijk} where

$$y_{ijk} = v_{ij(k)} - v_{ik(j)} = \boldsymbol{\mu}_{ij} - \boldsymbol{\mu}_{ik} + \epsilon_{ij(k)} - \epsilon_{ik(j)}. \tag{6}$$

It can be either assumed that judgments across choice sets are independent or dependent on the individual level. Under independence, one obtains

$$y_{ijk} = \boldsymbol{\mu}_{ij} - \boldsymbol{\mu}_{ik} + \epsilon_{ijk}, \tag{7}$$

where $\epsilon_{ijk} = \epsilon_{ij(k)} - \epsilon_{ik(j)}$ and the variability in the judgments comes solely from the ϵ's which are independently distributed given the person-specific parameters. For multiple comparisons of J options, the latent pairwise judgments of judge i can then be written conveniently as a linear

model. For example, in a comparison of the options j, k, l, and m, we obtain

$$
\boldsymbol{y}_i = \begin{pmatrix} y_{ijk} \\ y_{ijl} \\ y_{ijm} \\ y_{ikl} \\ y_{ikm} \\ y_{ilm} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\mu}_{ij} \\ \boldsymbol{\mu}_{ik} \\ \boldsymbol{\mu}_{il} \\ \boldsymbol{\mu}_{im} \end{pmatrix} + \begin{pmatrix} \epsilon_{ijk} \\ \epsilon_{ijl} \\ \epsilon_{ijm} \\ \epsilon_{ikl} \\ \epsilon_{ikm} \\ \epsilon_{ilm} \end{pmatrix} = A\boldsymbol{\mu}_i + \boldsymbol{\epsilon}_i, \qquad (8)
$$

where A is the paired-comparison design matrix.

There are important links between multivariate, non-linear mixed IRT models for nominal and binomial responses and random-utility models (Rijmen, Tuerlinckx, De Boeck, & Kuppens, 2003). Many of these models are based on random-effects versions of Luce's (1959) choice model which is consistent with a RUM model if and only if the random component in (2) follows a Gumbel distribution (Holman & Marley reported in Luce & Suppes, 1965). Examples include Bock's (1969, 1972) multinomial logit and nominal models, McFadden's (1974) conditional logit model, McFadden and Train's (2000) mixed multinomial logit model, Böckenholt's (2001b) ranking model, and Skrondal and Rabe-Hesketh's (2003) multilevel logit model.

The interpretation of the RUM model parameters is limited by the comparative and discrete nature of choice data. Most importantly, it is not possible to identify the origin and scale of the individual utility scales. One option may be preferred to another but this result does not allow any conclusions about the attractiveness level of the options or about interpersonal differences in the utilities (Guttman, 1946). Tsai (2000, 2003) discusses the identifiability of parameter estimates for Thurstonian ranking and paired comparison models (see also Tsai & Böckenholt, 2002). An important implication of this work is that the Case distinctions that were originally proposed by Thurstone (1927) are misleading and have to be interpreted as equivalence classes of covariance structures. However, as noted already by Thurstone and Jones (1957), it is possible to identify the origin of the utility scale by extending the choice task, for example, by including comparisons between pairs of options. Böckenholt (2004) provides a recent review and discussion of different methods that can be used for this purpose. In general, methods for identifying an utility scale origin are not only instrumental for avoiding difficulties in the interpretation of the estimated parameters of a choice model, they also provide useful insights about the underlying judgmental process.

Current Uses of Thurstonian Random Utility Models

This section discusses three major current psychometric research streams in the analysis of choice data. First, much work is underway to enhance our current tool set for modeling individual differences both at a particular point in time and over time. Second, there is strong interest in going beyond modeling the relationship between choice options and decision-makers and to consider a wider range of influences including social context and social influences. A third research stream focusses on ways to supplement and complement choice data because choice outcomes alone provide only limited information about the choice process and its determinants.

Individual Differences

The analysis of individual difference effects is one of most fruitful aspects of Thurstonian random utility models. Past work considered nested and crossed random effects, person-specific regression weights, and factor structures to capture dependencies among options under flexible distributional assumptions that include mixtures of the extreme value distribution, multivariate

normality, scale mixtures of multivariate normal distributions, and Dirchlet processes. For example, the following decomposition of an individual utility vector allows for the effects of judge-specific covariates x_i as well as observed and unobserved attributes of the options, given by M and Λ, respectively,

$$\mu_i = \mu + K x_i + M \beta_i + \Lambda f_i + v_i, \tag{9}$$

where K, β_i, and f_i contain regression coefficients. Effects that are not accounted for by the covariates are represented by v_i. Under normality assumptions, $v_i \sim N(0, \Sigma_v)$, $\beta_i \sim N(\beta, \Sigma_\beta)$, and $f_i \sim N(0, \Sigma_f)$. Although the interpretation of individual taste differences is much simplified under the assumption that decision-makers use the same set of attributes in assessing the choice options, some care must be taken in relating these attributes to the option utilities. For example, an ideal-point model may prove superior to the linear representation provided by the factor-analytic model in (9) if individuals choose the option that is closest to their "ideal" or most preferred option (Böckenholt, 1998; Coombs 1964; MacKay, Easley, & Zinnes, 1995). Applications of ranking and paired comparison models with unobserved attributes are reported by Brady (1989), Chan and Bentler (1998), Maydeu-Olivares and Böckenholt (2005), Takane (1987), Tsai and Böckenholt (2001), and Yu, Lam, and Lo (1998). Examining preferences for multi-attributed options, conjoint analysis (Bock & Jones, 1968; Marshall & Bradlow, 2002) uses (9) for the development of new products with optimal combinations of attributes.

When preferences for only a single choice set are elicited, the effects of between- and within-judge variability are confounded. For example, if judges are asked to rank J options only once, it is not possible to identify separately the effects of ϵ_{ij} and μ_{ij} in (2). Fortunately, with multiple choice sets or when choices are observed over time, it becomes possible to distinguish explicitly these sources of variability. Specifically, in the context of time-dependent choices, parameter-driven dynamic approaches have proven useful to capture dependencies on the within- and between-subject levels (Böckenholt, 2002; Allenby & Lenk, 1994). For example, in the context of modeling voting data collected before and after the 1992 presidential election, Böckenholt (2002) applied the following vector-autoregressive model with

$$\mu_{it} = \Upsilon \mu_{i,t-1} + \xi_{it}, \tag{10}$$

where μ_{it} are the options' mean evaluations of person i at time t, and the $(J \times J)$ coefficient matrix Υ contains the multivariate regression effects between the evaluations of the options at two adjacent time points. The diagonal elements of Υ contain each option's autoregressive term and the off-diagonal elements contain the lagged effects of the $J - 1$ remaining random utilities. The J-dimensional random effect ξ_{it} captures the part of μ_{it} that cannot be predicted on the basis of past evaluations of the options. There are many interesting variations on studying temporal effects on choice including time-continuous models choice (Böckenholt, 2005) and latent change models (Böckenholt & Dillon, 1997b).

Aside from temporal dependencies, there is also strong interest in modeling proximity–based dependencies that are introduced by known or unknown (latent) relationships among individuals (Anselin, 2002; Bradlow et al., 2005). This work goes beyond the standard assumption that individuals make choices in isolation by allowing explicitly for the possibility that choices of individuals are correlated or influenced by each other.

Social Dependencies

During the last decade the study of the effect of social interactions on choice has become an important area of research. Although it has long been recognized (e.g., Veblen, 1899; Leibenstein, 1950) that choices of individuals may depend on the choice behavior of their social network, empirical research on this topic has been slow, mainly, because of methodological problems

(Manski, 2000; Brock & Durlauf, 2001). Modeling issues that need to be addressed in empirical work on social interactions are:

(a) the identification of the reference group for which social interaction effects are sought to be established;
(b) self-selection processes of peer or group members;
(c) controls for correlated effects that affect all group members in a similar way; and
(d) controls for contextual effects such as exogenous social background characteristics of group members.

Because these issues are difficult to tackle in any observational study, progress is most likely to be made by combining observational with quasi-experimental or laboratory studies.

A basic version of an interaction-based model is obtained by extending (2) to include the influence of the choice behavior of others (Soetevent & Kooreman, 2007).

$$\upsilon_{ij} = \boldsymbol{\mu}_{ij} + \epsilon_{ij} + \lambda_{ij}, \tag{11}$$

where $\lambda_{ij} = \gamma w_{ij} \sum_{k \neq i} w_{kj}$ and $w_{ij} = 1$ if option j is chosen by person i and -1, otherwise. Thus, for $\gamma > 0$, the utility of choosing option j when another person chooses the same option as well is larger than the utility of choosing this option when it is not chosen by another person. For $\gamma = 0$, the model reduces to the standard Thurstonian choice model. Even for the simple model (11), with fixed effects for the social interaction component, multiple equilibria exist that complicate parameter estimation.

In general, the modeling of interaction structures has been based mainly on mathematical models originating in the area of statistical mechanics (Yeomans, 1992). Although the physical interpretation of these models may be of little interest to social scientists, their mathematical properties are intriguing and deserve to be explored in experiments on social decision-making.

Beyond Choice Data

Because choice data contain little information about the underlying choice process and its determinants, there have been many attempts both to supplement them by considering other data sources such as reaction times or process-tracing data (Böckenholt & Hynan, 1994; Johnson & Busemeyer, 2005) and to complement them by combining revealed and stated preference data (Ben-Akiva et al., 1997), or by collecting comparative judgment data on perceived attributes of the choice options. Below I discuss examples of both approaches by considering risky and multiattribute choices.

Uncertain Choice Outcomes

Many, if not most, real-life decisions are based on a mix of information and subjective expectations about the choice options under consideration. For example, the selection of a job may be based on a job description as well as on expectations about the career path. Purchases of over-the-counter drugs may be influenced by the drugs' ingredients but also by expectations about the ingredients' effectiveness and quality perceptions of the manufacturers. If these subjective expectations are rational or well-calibrated, it is possible to infer both expectations and utilities from choices alone. However, there is much evidence to suggest that this assumption is difficult to justify in general (see Kahneman, 2003, for a recent review). As a possible solution to this dilemma, Manski (2004) proposed to measure separately expectations in the form of subjective probabilities and combining them with the choice outcomes. Although much care needs to be taken in the elicitation of subjective probabilities, this approach is likely to mitigate problems arising from assuming that expectations are rational.

Consider option j with M outcomes $(e_{j1}, \ldots, e_{jM})$. By specifying the utility function to be additively separable in outcome and the outcomes to be binary, (2) can be extended to

$$v_{ij} = \sum_{m=1}^{M} \left[P_{ij}(e_m = 1)\mathbf{\mu}_{im(e_j=1)} + (1 - P_{ij}(e_m = 1))\mathbf{\mu}_{im(e_j=0)} \right] + \epsilon_{ij}, \tag{12}$$

where ϵ_{ij} denotes the unsystematic response variability in the evaluation of option j and $P_{ij}(e_m = 1)$ is the subjective probability that outcome m occurs when option j is selected. A simpler version of (12) was proposed by Shuford, Jones, and Bock (1960) who asked respondents to choose between lotteries with known outcome values and probabilities (see also Böckenholt, 2004). Although (12) is a useful starting point for studies on decision-making under uncertainty, it may require further extensions that take into account nonbinary outcomes and rank dependence in the valuation of the option outcomes and their subjective probabilities (Luce, 2000). However, this call for additional research should not distract from the observation that combining preference data with subjective probabilities is a promising avenue for psychometric analyses of decision-making when only partial information about the choice options is available.

Multiattribute Choices

In many applications, it is useful to collect choice data with respect to several attributes to understand how they relate to each other. In this case, one needs to distinguish between two sources of response dependencies introduced by the comparison of the different options with respect to the same attribute (within-attribute dependence), and by the evaluation of the same options on different attributes (between-attribute dependence). Böckenholt (1996) suggested the following decomposition of the covariance matrix of the multivariate judgments $(\mathbf{r}_{ia}, \mathbf{r}_{ib}, \ldots, \mathbf{r}_{iq})$ with respect to q attributes:

$$\mathbf{\Sigma_\mu} = \mathbf{\Sigma}_C \otimes \mathbf{\Sigma}_O,$$

where $\mathbf{\Sigma}_C$ represents the associations among the q attributes, and $\mathbf{\Sigma}_O$ represents the associations among the J options. For example, in the two-attribute case, when the covariance matrix of the options is represented by an identity matrix, we obtain

$$\mathbf{\Sigma_\mu} = (\rho_{ab}\mathbf{I}) \otimes \mathbf{I} = \begin{pmatrix} \mathbf{I} & \rho_{ab}\mathbf{I} \\ \rho_{ab}\mathbf{I} & \mathbf{I} \end{pmatrix},$$

and ρ_{ab} is the correlation between the attributes a and b.

Future Uses of Thurstonian-Based Analyses

Parallel to these expansions of random utility models, a growing literature has developed over the years that questions the descriptive value of RUMs. This work showed that the consideration of unsystematic error in (2) is insufficient to account for inconsistent choice behavior. Already May (1954) presented convincing evidence that strategies for comparing alternatives can lead to systematic violations of such internal consistency criteria as weak stochastic transitivity which is defined for the three options j, k, and l as

$$\Pr(j \succ k) \geq .50 \quad \text{and} \quad \Pr(k \succ l) \geq .50 \quad \Rightarrow \quad \Pr(j \succ l) \geq .50.$$

More generally, in the same way as the meaning of a word is clarified by the context of the sentence in which it appears, the valuation of a choice option seems to depend frequently on the

context of the choice set. The "asymmetric dominance" and the "compromise" effects are two well-known examples of systematic violations of expansion consistency (Simonson & Tversky, 1992). In the asymmetric dominance case, the dominance relation between two options affects their evaluations. This effect is not present when each of the options are presented in combination with an option that is neither dominating or dominated. The compromise effect occurs when a two-attribute option is presented as falling between two extreme options that are strong on one but weak on the other attribute. In this case, the "compromise" option seems to receive a boost for having medium values on both attributes. Such relational effects cannot be captured by models that consider context-independent evaluations of choice options.

Shafir and LeBoeuf (2002) review a long list of factors that have been shown to affect decision processes as well as possible explanations that can account for these effects. This list includes contextual effects (e.g., relational features such as dominance among choice options, temporal features of the choice situation), choice processes (e.g., the evaluability of options, decision strategies, information search strategies), presentation formats, frames, cultural and social norms, as well as characteristics of the decision-maker (e.g., emotional state, general intelligence, numeracy). In view of this list, it is perhaps surprising how well (2) can work in providing satisfactory accounts of choice data.

Because of the experimental nature of studies investigating systematic violations of RUMs, the literature contains only a few quantitative accounts on the degree to which observed choice behavior is consistent or inconsistent with a random utility representation (Ryan, Netten, Skatun, & Smith, 2006). This lack of information contributes to the ongoing controversy about the status of RUM models—whether they should be abandoned or extended in specific ways to increase their applicability. To resolve this debate, meta-analyses are needed but also new modeling approaches that allow identifying factors that characterize consistent and inconsistent choosers. For example, the π^* model (Rudas, Clogg, & Lindsay, 1994) could be employed usefully by including one mixture component based on (2) and a second one that captures deviations from this model to estimate the proportion of judges whose choices are inconsistent with a random utility representation. Alternatively, multiheuristic mixture models could be developed where each mixture component represents a different choice heuristic (Mellers & Biagini, 1994; Gonzalez-Vallejo, 2002; Wedel & Kamakura, 1999) that may lead to both consistent as well as inconsistent choices. A third avenue towards a more descriptive account of choice behavior is to replace the notion that utilities are predetermined and stable by a more dynamic view of utilities (Liechty, Fong, & DeSarbo, 2005). This approach was taken by Tsai and Böckenholt (2006) who introduced a more general framework for the analysis of pairwise judgments that does not require the two assumptions of independent judgments on the individual level and fixed utilities across choice sets. Instead, they decomposed the latent judgments as

$$v_{ij(k)} = \mu_{ij(k)} + \epsilon_{ij(k)},$$
$$v_{ik(j)} = \mu_{ik(j)} + \epsilon_{ik(j)},$$

where $\epsilon_{j(k)}$ and $\mu_{ij(k)}$ denote the random variation and person-specific evaluation of option j, respectively, when compared to option k. As a result, (6) changes to

$$y_{ijk} = v_{ij(k)} - v_{ik(j)} = \mu_{ij(k)} - \mu_{ik(j)} + \epsilon_{ij(k)} - \epsilon_{ik(j)}.$$

For multiple comparisons with J items, the paired comparison judgments can again be written conveniently as a linear model. Let $\boldsymbol{\epsilon} = (\epsilon_{1(2)}, \epsilon_{1(3)}, \ldots, \epsilon_{1(J)}, \epsilon_{2(1)}, \epsilon_{2(3)}, \ldots, \epsilon_{,(J)(J-1)})'$ be a $J(J-1)$-dimensional random vector, and let $\boldsymbol{B}$ be a $\binom{J}{2} \times J(J-1)$ matrix where each column of $\boldsymbol{B}$ corresponds to one of the $\binom{J}{2}$ paired comparisons and each row to one of the $J(J-1)$ terms

in ϵ_i. For example, when $J = 3$, $\epsilon = (\epsilon_{1(2)}, \epsilon_{1(3)}, \epsilon_{2(1)}, \epsilon_{2(3)}, \epsilon_{3(1)}, \epsilon_{3(2)})'$ and

$$
B = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix}. \tag{13}
$$

For example, in a comparison of the items j, k, l, and m with context-free mean evaluations of the choice options, we obtain

$$
y_i = \begin{pmatrix} y_{ijk} \\ y_{ijl} \\ y_{ijm} \\ y_{ikl} \\ y_{ikm} \\ y_{ilm} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \mu_{ij} \\ \mu_{ik} \\ \mu_{il} \\ \mu_{im} \end{pmatrix} + \begin{pmatrix} \epsilon_{ij(k)} - \epsilon_{ik(j)} \\ \epsilon_{ij(l)} - \epsilon_{il(j)} \\ \epsilon_{ij(m)} - \epsilon_{im(j)} \\ \epsilon_{ik(l)} - \epsilon_{il(k)} \\ \epsilon_{ik(m)} - \epsilon_{im(k)} \\ \epsilon_{il(m)} - \epsilon_{im(l)} \end{pmatrix} = A\mu_i + B\epsilon_i,
$$
$$\tag{14}$$

where $\epsilon_i = (\epsilon_{ij(k)}, \epsilon_{ij(l)}, \epsilon_{ij(m)}, \epsilon_{ik(j)}, \epsilon_{ik(l)}, \epsilon_{ik(m)}, \ldots, \epsilon_{im(j)}, \epsilon_{im(k)}, \epsilon_{im(l)})'$.

Decomposing $\mu_i = \mu + v_i$ and under multivariate normality assumptions for $\epsilon_i \sim N(0, \Sigma_\epsilon)$, $v_i \sim N(0, \Sigma_v)$, and ϵ_i independent of v_i, we obtain for the joint distribution of v_i, ϵ_i, and y_i,

$$
\begin{bmatrix} v_i \\ \epsilon_i \\ y_i \end{bmatrix} \sim N \left(\begin{bmatrix} 0_J \\ 0_{J(J-1)} \\ A\mu \end{bmatrix}, \begin{bmatrix} \Sigma_v & 0' & \Sigma_v A' \\ 0 & \Sigma_\epsilon & \Sigma_\epsilon B' \\ A\Sigma_v & B\Sigma_\epsilon & \Sigma_y \end{bmatrix} \right), \tag{15}
$$

and $\Sigma_y = A\Sigma_v A' + B\Sigma_\epsilon B'$.

The covariance matrix Σ_ϵ can provide useful insights about the dependencies between the options' utilities over different choice sets. For example, a high covariance between the utilities of option j across different option pairs indicates strong consistency in the choices involving this option and hence a low incidence of intransitive choices. A parsimonious, and easy to interpret, representation of Σ_ϵ is obtained when decomposing it as

$$
\Sigma_\epsilon = \sum_{j=1}^{J} (D_j \otimes S_j), \tag{16}
$$

where D_j is a $J \times J$ matrix with positive (j, j) element d_j and zero otherwise, and S_j is a $(J - 1) \times (J - 1)$ covariance matrix for option j.

Tversky's (1969) Gambling Study

Tsai and Böckenholt (2006) applied (16) in a replication of Tversky's (1969) intransitivity study. In this experiment, subjects were presented with four lotteries of the form (x, p; 0) where the amount x can be won with probability p and nothing otherwise. The winning probabilities of the four lotteries were $\frac{7}{24}$, $\frac{8}{24}$, $\frac{9}{24}$, and $\frac{10}{24}$ and the corresponding payoffs were $20, $19, $18, and $17, respectively. The gambles were constructed in such a way that payoffs are negatively correlated with expected values. Probabilities were presented in the form of difficult-to-compare pie sector diagrams, with the expectation that some participants would ignore small probability differences and focus on the payoffs instead. For large differences between the probabilities, it was expected that participants would consider probabilities in combination with the payoff in their decision to select the more attractive lottery. To test this hypothesis formally, Tsai and

Böckenholt (2006) specified the following covariance matrix for the lottery ($20, $\frac{7}{24}$; 0) when compared to the other three lotteries:

$$S_a = \begin{pmatrix} 1 & -\tau & -\tau \\ -\tau & 1 & \tau \\ -\tau & \tau & 1 \end{pmatrix}. \tag{17}$$

Because of the expected shift in the decision strategy, evaluations of lottery ($20, $\frac{7}{24}$; 0) when compared to lottery ($19, $\frac{8}{24}$; 0) were expected to be negatively correlated. In contrast, evaluations of lottery ($20, $\frac{7}{24}$; 0) when compared with lotteries ($18, $\frac{9}{24}$; 0) and ($17, $\frac{10}{24}$; 0) were expected to be positively correlated since they are based on similar judgmental processes in this case. The correlational pattern for the other three lotteries can be derived by applying the same reasoning.

The covariance structure (16) in combination with (17) was estimated for single subject data. Table 1 displays the observed binary and trinary marginal probabilities as well as the model predictions for one of the subjects and shows that (16) can capture well the dependencies between the repeated evaluations of the lotteries (with $\hat{\tau} = .63$ (.06)) and, equally important, it can account for the observed violations of weak stochastic transitivity. Thus, in contrast to the standard formulation of random utility models which assume independent judgments on the individual level and fixed utilities across choice sets, (16) can represent both consistent and inconsistent choices satisfactorily.

Tsai and Böckenholt (2007) illustrate further the usefulness of distinguishing explicitly between within- and between-judge variability based on (15) in combination with (16) in a reanalysis of the intertemporal choice data reported by Roelofsma and Read (2000). As in the replication of the Tversky (1969) study, (16) proved to be well-suited to account for systematic transitivity violations which were caused in this study by inconsistent trade-offs between "time" and "money" attributes characterizing the choice options. Extensions of the model framework to analyze other choice anomalies are currently underway. These studies show that both the compromise and the attraction effect can be modeled parsimoniously using the decomposition (16) of the within-judge covariance matrix (Böckenholt & Tsai, 2007).

Relaxing the assumption of fixed and predetermined utilities appears to be a promising approach to describe seemingly inconsistent choice behavior. By allowing the individual-level utilities to vary in repeated evaluations of the same options in different choice sets, this approach

TABLE 1.
Observed and expected probabilities for RUM models with stable and correlated utilities

	Obs.	Stable utilities	Corr. utilities
$\Pr(a \succ b)$	.61	.50	.63
$\Pr(a \succ c)$	.47	.46	.44
$\Pr(a \succ d)$	.43	.50	.42
$\Pr(b \succ c)$	.52	.45	.50
$\Pr(b \succ d)$	.38	.49	.43
$\Pr(c \succ d)$	.58	.54	.57
$\Pr(a \succ b, b \succ c, c \succ a)$	.11	.12	.10
$\Pr(b \succ a, c \succ b, a \succ c)$	.22	.12	.21
$\Pr(b \succ c, c \succ d, d \succ b)$	.11	.12	.12
$\Pr(c \succ b, d \succ c, b \succ d)$	.20	.12	.19

$a = ($20, $\frac{7}{24}$; 0)$; $b = ($19, $\frac{8}{24}$; 0)$; $c = ($18, $\frac{9}{24}$; 0)$; $d = ($17, $\frac{10}{24}$; 0)$.

can accomodate a considerably wider range of inconsistent choice behavior than has been possible so far. Equally important, with this framework it becomes possible to relate the utilities' reliability estimates to both context- and person-specific covariates with the result that one can test rigorously determinants of variables and possibly inconsistent utility assessments.

Conclusion

Do people choose the options they enjoy most? There is much evidence to suggest that the answer is negative: People do not always know what they like and their ability to forecast future utilities of potential choices appears to be systematically biased (Kahneman & Thaler, 2006). The subsequent question about necessary and/or sufficient (environmental) conditions that facilitate utility maximization has received less attention so far. Notable exceptions include the notion of "libertarian paternalism" (Sunstein & Thaler, 2003) which sets up default options in such a way as to help people in their utility maximization. Along the same lines, a recent large-scale study by McFadden (2006) on choices among Medicare-approved plans demonstrates the importance of aiding consumers in helping themselves instead of relying on their self-interest in making optimal choices (see also Lynch & Wood, 2006).

Although behavioral research on choice behavior points convincingly to limitations of random utility models, few would dispute the usefulness of these models in rendering parsimonious descriptions of how individuals perceive and evaluate choice options. Because they are based on a limited explanatory framework for how people make choices, random utility models can provide both a parsimonious quantitative description of choice outcomes and a flexible framework for modeling individual and contextual differences in choice behavior at a particular point in time and over time. However, generalizations of RUM results to different choice situations and options require care and need to be based on additional validation studies. The consideration of random response errors and unstable preferences alone is not sufficient to account for deviations from utility maximization.

There are many research opportunities on the horizon for psychometricians. Better measures and indicators of utility are needed that capture the hedonic experience associated with a choice outcome. To infer utility from choices alone without taking into account anticipated or experienced hedonic reactions no longer seems sufficient (Kahneman, 2003). Measures of brain activity and brain electrochemistry in combination with experimental treatments are starting to become available to provide much needed insights on the links between choice and sensations of pleasure and pain but statistical methods are lacking for effective analyses of these links (Montague, King-Casas, & Cohen, 2006). The development of new measures and indicators is facilitated further by expanding their connections to psychological concepts and processes. For example, affective and motivational mechanisms are becoming integrated in theories of individual choice behavior and have led to new concepts such as "irrational wanting" and "subrational liking" (Winkielman & Berridge, 2003), pointing to obvious limitations in current choice modeling frameworks. Recently, Mourali, Böckenholt, and Laroche (in press) showed that the compromise and asymmetric dominance effects can be weakened or strengthened depending on whether a person has a promotion or prevention focus (Higgins, 1997), demonstrating that motivational factors need to be taken into account when modeling choice data. In general, the search for choice models that are behaviorally more realistic but still tractable is complicated greatly by identifiability and endogeneity issues. Current choice models for unstable preferences, uncertain outcomes, or social interactions suffer from difficulties both in separating different sources of heterogeneity and in identifying multiple equilibria, all of which needs to be studied carefully with the help of well-designed empirical studies. Clearly, challenges for modeling and predicting choice behavior are abundant but they also assure the future well-being of psychometrics.

References

Allenby, G.M., & Lenk, P.J. (1994). Modeling household purchase behavior with logistic normal regression. *Journal of the American Statistical Association*, *89,* 1218–1231.

Anselin, L. (2002). Under the Hood: Issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, *17,* 247–267.

Ansari, A., & Iyengar, R. (in press). Semiparametric Thurstonian models for recurrent choices: A Bayesian analyis. *Psychometrika*, *71.*

Ashford, J.R., & Sowdon, R.R. (1970). Multivariate probit analysis. *Biometrics*, *26,* 535–546.

Ben-Akiva, M., McFadden, D., Abe, M., Böckenholt, U., Bolduc, D., Gopinath, D., Morikawa, T., Ramaswamy, V., Rao, V., Revelt, D., & Steinberg, D. (1997). Modeling methods for discrete choice analysis. *Marketing Letters*, *8,* 273–286.

Block, H., & Marschak, J. (1960). Random orderings and stochastic theories of response. In I. Olkin et al. (Eds.), *Contributions to probability and statistics* (pp. 97–132). Stanford, CA: Stanford University Press.

Bock, R.D. (1958). Remarks on the test of significance for the method of paired comparisons. *Psychometrika*, *23,* 323–334.

Bock, R.D. (1969). Estimating multinomial response relations. In R.C. Bose et al. (Eds.), *Essays in probability and statistics* (pp. 111–132). Chapel Hill: University of North Carolina Press.

Bock, R.D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, *37,* 29–51.

Bock, R.D., & Jones, L.V. (1968). *The measurement and prediction of judgment and choice*. San Francisco: Holden-Day.

Böckenholt, U. (1992). Thurstonian models for partial ranking data. *British Journal of Mathematical and Statistical Psychology*, *45,* 31–49.

Böckenholt, U. (1996). Analyzing multi-attribute ranking data: Joint and conditional approaches. *British Journal of Mathematical and Statistical Psychology*, *49,* 57–78.

Böckenholt, U. (1998). Modeling time-dependent preferences: Drifts in ideal points. In M. Greenacre & J. Blasius (Eds.), *Visualization of categorical data* (pp. 461–476). Hillsdale, NJ: Erlbaum.

Böckenholt, U. (2001a). Thresholds and intransitivities in pairwise judgments: A multilevel analysis. *Journal of Educational and Behavioral Statistics*, *26,* 269–282.

Böckenholt, U. (2001b). Mixed-effects analyses of rank-ordered data. *Psychometrika*, *66,* 45–62.

Böckenholt, U. (2002). A Thurstonian analysis of preference change. *Journal of Mathematical Psychology*, *46,* 300–314.

Böckenholt, U. (2004). Comparative judgments as an alternative to ratings: Identifying the scale origin. *Psychological Methods*, *9,* 453–465.

Böckenholt, U. (2005). A latent Markov model for the analysis of longitudinal data collected in continuous time: States, durations, and transitions. *Psychological Methods*, *10,* 65–83.

Böckenholt, U., & Dillon, W.R. (1997a). Modeling within-subject dependencies in ordinal paired comparison data. *Psychometrika*, *62,* 414–434.

Böckenholt, U., & Dillon, W.R. (1997b). Some new methods for an old problem: Modeling preference changes and competitive market structures in pre-test market data. *Journal of Marketing Research*, *34,* 130–142.

Böckenholt, U., & Hynan, L. (1994). Caveats on a process-tracing measure and a remedy. *Journal of Behavioral Decision Making*, *7,* 103–118.

Böckenholt, U., & Tsai, R. (2006). Random-effects models for preference data. In C.R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (Vol. 26, pp. 447–468). Amsterdam: Elsevier Science.

Böckenholt, U., & Tsai, R. (2007). An unstable preference view of the compromise and asymmetric dominance effects. Manuscript in preparation.

Bradlow, E., Bronnenberg, B., Russell, G., Arora, N., Bell, D.R., Dev Suvvuri, S., Ter Hofstede, F., Sismeiro, C., Thomadsen, R., & Yang, S. (2005). Spatial models in marketing. *Marketing Letters*, *16,* 267–278.

Brady, H.E. (1989). Factor and ideal point analysis for interpersonally incomparable data. *Psychometrika*, *54,* 181–202.

Brock, W.A., & Durlauf, S.N. (2001). Interactions-based models. In *Handbook of econometrics* (Vol. 5, pp. 3297–3380). Amsterdam: North-Holland.

Caffo, B., & Griswold, M. (2006). A user-friendly tutorial on link-probit-normal models. *The American Statistician*, *60,* 139–145.

Chakravarti, D., Sinha, A., & Kim, J. (2005). Choice research: A wealth of perspectives. *Marketing Letters*, *16,* 173–182.

Chan, W., & Bentler, P.M. (1998). Covariance structure analysis of ordinal ipsative data. *Psychometrika*, *63,* 369–399.

Coombs, C.H. (1964). *A theory of data*. New York: Wiley.

David, H.A. (1988). *The method of paired comparisons*. London: Griffin.

Edgeworth, F.Y. (1881). *Mathematical physics*. London: Kegan Paul.

Falmagne, J.C. (1985). *Elements of psychophysical theory.* Oxford: Clarendon Press.

Fechner, G.T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf und Härtel.

Georgescu-Roegen, N. (1936). The pure theory of consumer's behavior. *Quarterly Journal of Economics*, *50,* 545–593.

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

Gueorguieva, R., & Agresti, R. (2001). A correlated probit model for joint modeling of clustered binary and continuous responses. *Journal of the American Statistical Association*, *96,* 1102–1112.

Guttman, L. (1946). An approach for quantifying paired comparisons and rank order. *Annals of Mathematical Statistics*, *17,* 144–163.

Hanemann, M. (1984). Discrete continuous models of consumer demand. *Econometrica*, *52,* 541–561.

Hausman, J., & Wise, D. (1978). A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences. *Econometrica*, *46,* 403–429.

Hensher, D.A., Rose, J.M., & Greene, W.H. (2005). *Applied choice analysis: A primer*. Cambridge: Cambridge University Press.

Higgins, E.T. (1997). Beyond pleasure and pain. *American Psychologist*, *52,* 1280–1300.

Johnson, J.G., & Busemeyer, J.R. (2005). A dynamic, stochastic, computational model of preference reversal phenomena. *Psychological Review*, *112,* 841–861.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93,* 1449–1475.

Kahneman, D., & Thaler, R.H. (2006). Anomalies: Utility maximization and experienced utilty. *Journal of Economic Perspectives*, *20,* 221–234.

Leibenstein, H. (1950). Bandwagon, snob, and Veblen effects in the theory of consumerŠs demand. *The Quarterly Journal of Economics*, *64,* 183–207.

Liechty, J., Fong, D.K.H., & DeSarbo, W.S. (2005). Dynamic models incorporating individual heterogeneity: Utility evolution in conjoint analysis. *Marketing Science*, *24,* 285–293.

Louviere, J.J., Hensher, D.A. & Swait, J.D. (2000). *Stated choice methods*. New York: Cambridge University Press.

Luce, R.D. (1959). *Individual choice behavior*. New York: Wiley.

Luce, R.D. (2000). *Utility of gains and losses: Measurement-theoretical and experimental approaches*. Hillsdale, NJ: Erlbaum.

Luce, R.D., & Suppes, P. (1965). Preference, utility, and subjective probability. In R.D. Luce, R.R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. III, pp. 235–406). New York: Wiley.

Luce, R.D., & Tukey, J. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, *1,* 1–27.

Lynch, J.G., & Wood, W. (2006). Helping consumers help themselves. *Journal of Public Policy and Marketing*, *26,* 1–7.

MacKay, D.B., Easley, R.F., & Zinnes, J.L. (1995). A single ideal point model for market structure analysis. *Journal of Marketing Research*, *32,* 433–443.

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

Manski, C.F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives*, *14,* 115–136.

Manski, C.F. (2004). Measuring expectations. *Econometrica*, *72,* 1329–1376.

Marschak, J. (1960). Binary choice constraints on random utility indictor. In K.I. Arrow, S. Karlin, & P. Suppes (Eds.), *Stanford symposium on mathematical methods in the social sciences* (pp. 312–329). Stanford, CA: Stanford University Press.

Marshall, P., & Bradlow, E.T. (2002). A unified approach to conjoint analysis models. *Journal of the American Statistical Association*, *97,* 674–682.

May, K.O. (1954). Intransitivity, utility, and the aggregation of preference patterns. *Econometrica*, *22,* 1–13.

Maydeu-Olivares, A., & Böckenholt, U. (2005). Structural equation modeling of paired comparison and ranking data. *Psychological Methods*, *10,* 285–304.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). New York: Academic Press.

McFadden, D. (1984). Qualitative choice models. In Z. Griliches & M.D. Intriligator (Eds.), *Handbook of econometrics* (pp. 1395–1457). Cambridge, MA: MIT Press.

McFadden, D. (2001). Economic choices. *American Economic Review*, *91,* 351–378.

McFadden, D. (2006). Free markets and fettered consumers. *American Economic Review*, *96,* 5–29.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, *15,* 447-470.

Mellers, B.A., & Biagini, K. (1994). Similarity and choice. *Psychological Review*, *101,* 505–518.

Montague, P.R., King-Casas, B., & Cohen, J.D. (2006). Imaging valuation models in human choice. *Annual Review of Neuroscience*, *29,* 417–448.

Mourali, M., Böckenholt, U., & Laroche, M. (in press). Compromise and attraction effects under prevention and promotion motivations. *Journal of Consumer Research*.

Pareto, V. (1900). Sunto di alcuni capitoli di un nuovo trattato di economia pura del Prof. Pareto. *Giornale degli Economisti*, *20,* 216–235, 511–549.

Rijmen, F., Tuerlinckx, F., De Boeck, P., & Kuppens, P. (2003). A nonlinear mixed model framework for item response theory. *Psychological Methods*, *8,* 185–205.

Roelofsma, P., & Read, D. (2000). Intransitive intertemporal choice. *Journal of Behavioral Decision Making*, *13,* 161–177.

Rudas, T., Clogg, C.C., & Lindsay, B.G. (1994). A new index of fit based on mixture methods for the analysis of contingency tables. *Journal of the Royal Statistical Society, Series B*, *56,* 623–639.

Ryan, M., Netten, A., Skatun, D., & Smith, P. (2006). Using discrete choice experiments to estimate a preference-based measure of outcome—An application to social care for older people. *Journal of Health Economics*, *25,* 927–944.

Shafir, E., & LeBoeuf, R.A. (2002). Rationality. *Annual Review of Psychology*, *53,* 491–517.

Shuford, E.H., Jones, L.V., & Bock, R.D. (1960). A rational origin obtained by the method of contingent paired comparisons. *Psychometrika*, *25,* 343–356.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, *29,* 281–295.

Skrondal, A., & Rabe-Hesketh, S. (2003). Multilevel logistic regression for polytomous data and rankings. *Psychometrika*, *68,* 267–287.

Soetevent, A.R., & Kooreman, P. (in press). A discrete choice model with social interactions: With an application to high school teen behavior. *Journal of Applied Econometrics*.

Sunstein, C.R., & Thaler, R.H. (2003). Libertarian paternalism is not an oxymoron. *University of Chicago Law Review*, *70,* 1159–1202.

Takane, Y. (1987). Analysis of covariance structures and probabilistic binary choice data. *Cognition and Communication*, *20,* 45–62.

Thurstone, L.L. (1927). A law of comparative judgment. *Psychological Review*, *34,* 273–286.

Thurstone, L.L. (1931). The indifference function. *Journal of Social Psychology*, *2,* 139–167.

Thurstone, L.L., & Jones, L.V. (1957). The rational origin for measuring subjective values. *Journal of the American Statistical Association*, *52,* 458–471.

Train, K. (2003). *Discrete choice methods with simulations*. Cambridge, MA: MIT Press.

Tsai, R. (2000). Remarks on the identifiability of Thurstonian ranking models: Case V, Case III, or neither? *Psychometrika*, *65,* 233–240.

Tsai, R. (2003). Remarks on the identifiability of Thurstonian paired comparison models under multiple judgment. *Psychometrika*, *68,* 361–372.

Tsai, R., & Böckenholt, U. (2001). Maximum likelihood estimation of factor and ideal point models for paired comparison data. *Journal of Mathematical Psychology*, *45,* 795–811.

Tsai, R., & Böckenholt, U. (2002). Two-level linear paired comparison models: Estimation and identifiability issues. *Mathematical Social Sciences*, *43,* 429–449.

Tsai, R., & Böckenholt, U. (2006). Modeling intransitive preferences: A random-effects approach. *Journal of Mathematical Psychology*, *50,* 1–14.

Tsai, R., & Böckenholt, U. (2007). On the importance of distinguishing between within- and between-subject effects in intransitive intertemporal choice. Manuscript submitted for publication.

Tversky, A. (1969). Intransitivity of preference. *Psychological Review*, *76,* 31–48.

Veblen, T. (1899). *The theory of the leisure class*. New York: Macmillan.

Wallis, W.A., & Friedman, M. (1942). The empirical derivation of indifference functions. In O. Lange (Ed.), *Studies in mathematical economics and econometrics* (pp. 175–189). Chicago: University of Chicago Press.

Wedel, M., & Kamakura, W.A. (1999). *Market segmentation: Conceptual and methodological foundations*. Dodrecht: Kluwer Academic.

Winkielman, P., & Berridge, K. (2003). Irrational wanting and subrational liking: How rudimentary affective and motivational processes shape preferences and choice. *Political Psychology*, *24,* 657–680.

Yao, G., & Böckenholt, U. (1999). Bayesian estimation of Thurstonian ranking models based on the Gibbs sampler. *British Journal of Mathematical and Statistical Psychology*, *52,* 79–92.

Yeomans, J. (1992). *Statistical mechanics of phase transitions*. Oxford: Oxford University Press.

Yu, P.L.H., Lam, K.F., & Lo, S.M. (1998). Factor analysis for ranking data. Unpublished manuscript. Department of Statistics, The University of Hong Kong.