Formal Choice Models of Informal Choices:

What Choice Modeling Research Can (and Can't) learn from Behavioral Theory

Jordan J. Louviere

Robert J. Meyer*

* The authors are listed alphabetically. Jordan J. Louviere is Professor of Marketing and acting Executive Director Centre for the Study of Choice (CenSoC), Faculty of Business, University of Technology, Sydney, Australia. Robert J. Meyer is the Gayfryd Steinberg Professor Marketing and co-director of the Center for Risk and Decision Processes at the Wharton School of the University of Pennsylvania, Philadelphia, Pennsylvania.

Formal Choice Models of Informal Choices:

What Choice Modeling Research Can (and Can't) learn from Behavioral Theory

Abstract

In this paper we illustrate the benefits of forging a better alliance among behavioral, economic and statistical approaches to modeling consumer choice behavior. We focus on the problems that arise when building descriptive models of choice in evolving markets, where consumers are likely to have poorly-developed preferences and be influenced by beliefs about future market changes. We illustrate how understanding the actual process that is driving preferences can provide analysts with both better *a priori* insights into the model structures that are likely to provide the best descriptive account of choices in such settings, as well as how stable these structures are likely to be over time. We show, for example, that analogical reasoning heuristics—a common strategy for making decisions under preference uncertainty—can produce choice patterns that resemble the output of complex nonlinear, non-additive, multi-attribute utility rules. Likewise, because novice consumers are likely to display strong individual differences in the variance of unobserved components of utility, methods that fail to recognize such differences will tend to overstate the actual extent of taste heterogeneity that exists in a population. We also illustrate the benefits of a reverse dialogue, how economic theory can lead behavioral researchers to more parsimonious explanations for apparent anomalies in choice tasks where preferences are uncertain. We show, for example, that some *ad-hoc* models that have been used to statistically describe the compromise effect in choice can be deduced from first principles of rational risky decision making.

Introduction

Choice modeling research in marketing has evolved through the interplay of three different approaches to the study of human decision making. One is an economic perspective by which consumers make choices in a manner that is consistent with random utility maximization. Consumers are viewed as having well-developed preference functions defined over product attributes, and they choose those options whose attributes offer the most attractive tradeoffs either at the time of choice or in the short or long run (e.g., McFadden 1981). A second approach is exemplified by behavioral researchers and psychologists, who argue that actual choice processes may be far removed from the rational mechanisms assumed by economists. That is, to the extent that preferences exist at all, they discontinuous and imprecise, with choices being the outcome of heuristic rules that are uniquely constructed in response to the external appearance of options in choice sets (e.g., Payne, Betman, and Johnson 1993). A third and final view has grown rapidly since the late 1980s, which is a statistical approach to modeling choices. Adherents of this approach claim ideological neutrality in the debate over preferences and processes. That is, choices are simply viewed as data; any model of choice is fair game as long as it passes tests of descriptive and predictive validity in a given context (e.g., Abe 1996; ter Hofstede, Kim and Wedel 2002; Rossi, Allenby and McCulloch 2005; Kamakura and Wedel 2004). Not surprisingly, the statistical paradigm is less concerned with whether any given model can be deduced from first principles of utility maximization or cognitive theory.

Although the three approaches differ philosophically and to some extent methodologically, intuition suggests that that they might usefully converge over time as our understanding of choice behavior evolves and progresses. Yet, the academic reality appears to be very different. For example, behavioral researchers have tended to focus on laboratory demonstrations of failures of the assumptions of standard economic models (e.g., context invariance) and given limited attention to developing alternative modeling paradigms that might account for these failures (for exceptions, see, e.g., Kivetz, Oded, and Srinivasan 2004; Tversky and Simonson 1993). Adherents of the economics view of choice modeling, for their part, have frequently been dismissive of behavioral findings, arguing that lab settings exaggerate the size of errors that would be observed in real markets or that they can be captured through complex generalizations of standard models (e.g., Machina 1982). Finally, statistical modelers have done little to resolve theoretical gaps between the behavioral and economic camps. While there is much to be learned and gained from incorporating statistical advances from discrete multivariate and Bayesian statistics in choice modeling (e.g., Rossi and Allenby 2003), there is also much to be lost by adopting a purely statistical view of what is inherently a human behavioral process.

The purpose of this paper is to take a small step toward fusing these different perspectives in the analysis of choice data. We take a limited first step by exploring one dimension of this fusion, namely what empirical economic and statistical modelers can constructively learn from behavioral researchers when building models of consumer choice in evolving markets—a setting where applications of traditional methods have often been seem as problenmatic.

Choice and Market Evolution

Let us begin by with a thought experiment that illustrates the types of modeling challenges that we try to address in this paper. Consider a simple market with one (monopoly) provider of a good, such as a monopoly provider of cable or broadband services or another public utility market. In this market consumers must decide whether or not to choose the good. At some point the good in question is launched into the market, and we assume that prior to launch information is available about the good, its features and price(s) and the likely launch date. Thus, prior to launch some consumers in the market are aware that the service will be provided, and have reasonably complete information about its features and likely prices. Another, probably much larger, proportion of consumers is "vaguely aware" that the service will be provided, and has incomplete information about features and possible prices. Finally, a third proportion is unaware of the good or that it will be launched.

This market is thus characterized by three stylized groups of consumers, who can be viewed as being on a continuum of being aware and informed about the good, or they can be viewed as three discrete segments. Initially, the most aware and informed are likely to be a small minority; the vaguely aware and informed, while probably a larger proportion, also are likely to be a minority; with most of the market more likely to be unaware and uninformed. Then, the good is launched, and things begin to change. To the extent that the good is of interest to consumers and they are capable of buying and consuming it, which allows them to receive the associated benefits or problem solutions the good provides, we expect the proportion of consumers who are aware and informed to grow over time. Likewise, consumers who are unaware and uninformed gradually will move into the vaguely aware and informed group, and in this way the market will evolve from the "bottom up".

A marketer who wishes to model the decision of whether consumers choose the good in this market would typically begin with the tools of random utility theory. Each consumer n in the population would be assumed to associate with the new service i a utility $U_{nit} = \beta_n' X_{it} + \varepsilon_{nit}$, where $X_{it,}$ is a vector of the measured attributes of the service (e.g., price), β_n is an associated parameter vector describing the consumer's tastes for these attributes, and ε_{nit} is an unobserved component of utility. The unobserved component ε_{nit} would typically be assumed to follow an independently and identically distributed extreme value distribution (over consumers, choice alternatives, and service characteristics). If this assumption is satisfied, the individual choice probabilities can be represented by the well known multinomial logit model

$$P_{nit} = \frac{e^{\beta_n' X_{it}}}{\sum_k e^{\beta_n' X_{kt}} + \theta_{nt}}. \qquad (1)$$

where θ_{nt} is the consumer's utility for unmeasured outside goods. To extend expression (1) to the study of population or market choices, analysts typically make assumptions about how tastes β_n vary over the population. For example, if β_n can be assumed to have a stationary parametric distribution, then population choice can be modeled by a random-coefficients or mixed logit model that assumes $U_{nit} = \beta' X_{it} + \eta_{nit} + \varepsilon_{nit}$, where η_{nit} is a random variable that captures unobserved individual departures from a common strict utility $\beta' X$ (e.g., Hensher and Greene 2003).

It should be clear that while the above approach might provide a good statistical description of the association that exists between choices and service attributes at a particular point in time (or over a series of points in time) for a particular sample of people during the course of market evolution, it captures few of the behavioral features of service-choice dynamics mentioned above. For example, it does not characterize how parameter heterogeneity might be associated with factors that underlie differential levels of awareness and information possessed by consumers, the provider's decisions about communications and access, and beliefs held by consumers about the market's future (e.g., the possibility of new entrants or expectations about how the technology will evolve). While analysts may acknowledge that these associations are likely to change as a market evolves, exactly how the changes will occur or their trajectories typically lies outside the purview of the analysis. Thus, it is fair to say that the overwhelming majority of these types of models are purely descriptive with little real explanatory capability.

In the sections below we try to illustrate more precisely how real behavioral processes underlying choices in markets can manifest themselves in the data that are used to estimate reduced-form statistical models, and how overlooking behavioral processes may lead analysts to erroneous conclusions about both the nature of preferences in markets and how markets will evolve over time. More specifically, we explore what behavioral theory would predict about the empirical appearance and stability of typically-estimated random utility models when they are used describe the choice behavior of consumers who:

1. Have high levels of uncertainty about their preferences for goods in a market;
2. Use heuristic short-cuts that do not utilize all the product-attribute information available to them at the time of choice; and
3. Have strategic foresight; that is, consider how the current choice will affect the utility gained from future choices.

Modeling Choice by Naïve Consumers

One aspect of random utility theory that elicits few quarrels is the assumption that consumers are guided by a desire to choose the option that will give them the most utility or pleasure. But for many consumers, particularly those in newly evolving markets, the ability to achieve this maximization goal is inhibited by the simple fact that preferences are uncertain. For example, if a novice consumer were forced to decide whether it was worth $10 a month to adopt a broadband service that would increase download speeds by 100 kilobytes, the axioms of utility theory would not help her much to make this decision. To be useful, she would need to know what a kilobyte is, the amount of additional pleasure that she could expect from a 100kb, increase, and how to exchange the extra pleasure for dollars, knowledge few novice consumers are likely to have. The choices we observe in new markets, therefore, reflect an ambiguous mix of enduring preferences and the heuristics consumers use to overcome the *lack* of preferences.

What are the heuristics consumers use to overcome a lack of attribute preference knowledge? The consensus view is that naïve choices are often made using analogical reasoning. That is, when a new product is encountered, consumers judge it by recalling products that they consumed in the past with similar attributes (e.g., Gregan-Paxton and Roedder John 1997; Norman 1988). For example, this sort of pattern-matching process is thought to explain how people with little skill in mathematics can learn to play complex equilibria in games. Instead of encoding and solving optimal strategies, most games allow players to discover equilibria simply by being willing to repeat the moves that yielded the highest payoffs in the past (e.g., Camerer and Ho 1999; Fudenberg and Levine 1998). Similarly, the expertise of wine connoisseurs probably lies less in their skills at using algebraic rules to predict quality and more in their possession of a rich memory bank of past referent examples that form the basis of evaluations (known as smell and taste memory).

The pervasive use of pattern-matching heuristics in novel product judgments was illustrated in work by Meyer (1987), who examined the process by which consumers learn to make multi-attribute judgments in a novel product category. In his experiments participants were shown a series of product profiles described by several unfamiliar attributes and levels (copper alloys generated by different production methods), and they were asked to predict each product's likely quality (strength). Once the subjects made a prediction, they received feedback about the "true" quality of an option. Consistent with behaviors previously observed in tasks like this (e.g., Mellers 1981), after several rounds of feedback participants became quite good at making forecasts. That is, they acted "as if" they had learned the multi-attribute rule that determined true quality, and were using it to make forecasts. Yet, the surprising empirical outcome associated with this experiment was that in the course of making judgments a subset of respondents asked to provide concurrent verbal protocols gave no indication that they actually made judgments using a multi-attribute rule. Instead, they appeared to make their judgments using a pattern-matching process whereby they made forecasts about how similar any given new profile was to a previously-seen profiles with known qualities. Their judgments improved over time not because they were developing better knowledge of a rule, but because the database of referent examples improved, and this data base allowed them to effectively mimic the outcomes of such a rule.

Should choice modelers be worried by this result? Not necessarily; in the same way that a game theorist would be indifferent to whether people play equilibria because they actually calculate the optimal strategy or they stumble their way there by trial and error (Fudenberg and Levine 1998), one presumes that random utility theorists would be happy to view multi-attribute utility theory as an *as-if* model of the way in which people evaluate options. If pattern-matching heuristics produce judgment data that are well-approximated by stable linear-additive models, then one clearly can build a paradigm around this; mathematical convenience in this case would trump behavioral realism. But *is* there such an isomorphism, and how widespread might it be?

It is easy to show that a mathematical equivalence exists between pattern-matching and linear-additive rules, but only under a limiting condition of product-class experience: when the underlying (or true) reward structure is linear-additive in attributes and decision makers have had direct experience with all of the product profiles under study in a given multi-attribute space. In contrast, the more limited a person's judgmental experience in any given context, the less linear-additivity will describe their judgments, even if the underlying reward function is linear-additive.

To demonstrate, consider the following pattern-matching model of expected valuations:

$$EV_i = \delta_{iz}(V_z) + (1 - \delta_{iz})p \qquad (2)$$

Where EV_i is the anticipated utility of some multi-attribute profile *i*, δ_{iz} is a 0-1 bounded measure of the subjective similarity between profile *i* and the experienced profile *z* that is most similar to *i*, V_z is the experienced utility of profile $_z$, and p is a judgmental prior[1]. Suppose, also for the sake of simplicity, that *experienced* utility is given by the linear-additive multi-attribute model

$$V_k = \beta' X_k \tag{3}$$

where X_k is a vector of measures of the values attributes of profile *k* that is linear in *V*, and β is an associated weight vector.

It should be obvious that expression (2) is equivalent to (3) when δ_{iz} equals 1 for all comparisons *i,z*; that is, when every profile has been experienced by the consumer. In this case it truly doesn't matter whether consumers literally calculate utilities to form projective judgments (expression 2) or merely act *as if* they do (expression 1). But what about a more realistic case where consumers have limited experiences in a market; that is, for any given possible complete factorial array of profiles, what happens if only a small subset have valuations? In such cases (2) and (3) will not be equivalent; instead consumers will reveal an "as if" multi-attribute judgment rule that departs from the asymptotic (or full-information) rule (3) in predictable ways. Specifically, if we regress projected preferences (EV_i) against the attributes of each option we do not recover a true or asymptotic utility function (2), but instead a multiattribute function whose form is distorted by the similarity of each *i* to the actually experienced profiles *z*, their true valuations, and the judgmental prior; that is, $EV_i = \delta_{iz}(\beta' X_z) + (1 - \delta_{iz})p$.

To illustrate the properties of such an approximation, consider the case of a simple judgment context in which consumers evaluate the attractiveness of each of a

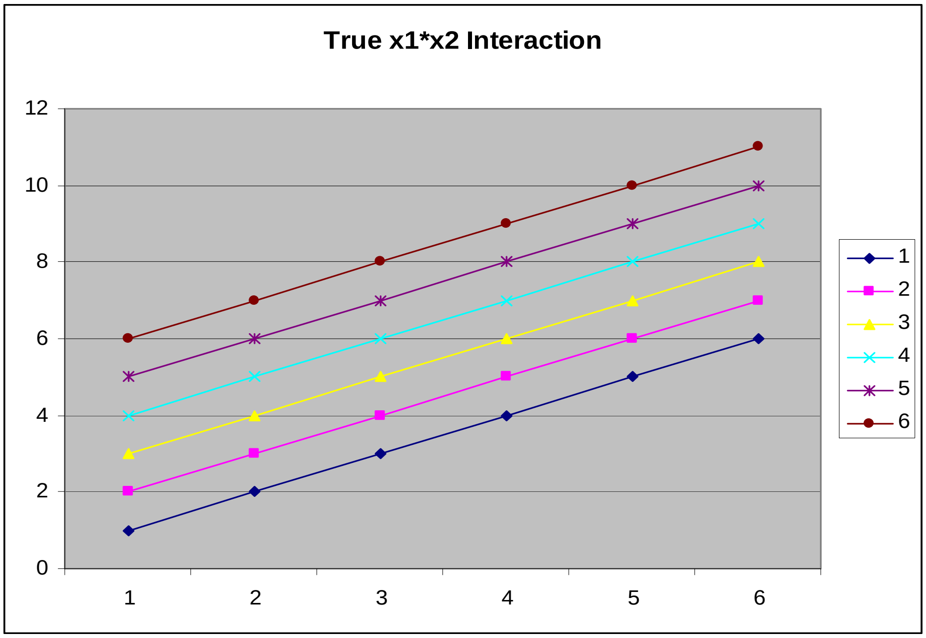
---

[1] Expression (2) can be seen as a special case of the general class of case-based decision models described by Gilboa amnd Schmeidler (1995) where the similarity function is defined over the maximum-similarity referent stored in memory.

number of two-attribute options on a subjective scale. Each option is described by a pair of levels that represent a combination of two six-level attributes. Respondents judge all 36 combinations represented by the factorial array. As above, for simplicity we assume that the true utility that would be observed by a consumer consuming each of the profiles is given by the additive rule $V_i = x_{i1} + x_{i2}$. Upon entering the task, however, a consumer has only limited experience in the category, and has directly experienced only a small subset of the 36 profiles. We consider the implications of approximating a pattern-matching judgment rule (expression 1) with a linear model in two illustrative cases: 1) the consumer's previous experiences correspond to the two extremes of the attribute space (the 1,1 and 6,6 profiles, respectively), and 2) the consumer has experienced three mid-valued options: the 2,2, 3,3, and 4,4 options. To generate a numeric example we used a normalized similarity metric $\delta_{iz} = \underset{k}{MAX}[1-(\sum_z | x_{iz} - x_{kz} | / \sum_z | x_{\max(z)} - x_{\min(z)})]^{\lambda}$ where λ=4, and assumed a prior (p) equal to the 1-12 response scale mean.

In Figure 2 we plot the resulting two-factor interaction graphs for each of these two cases (2b and 2c) to be contrasted with the normative interaction (2a).

Figure 2a



\

Figure 2b

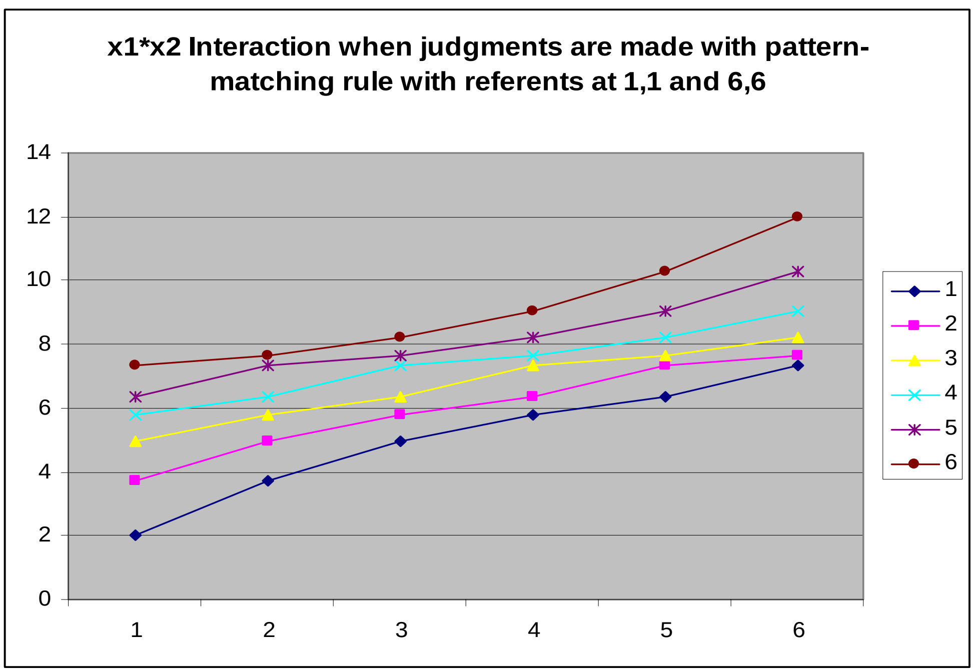
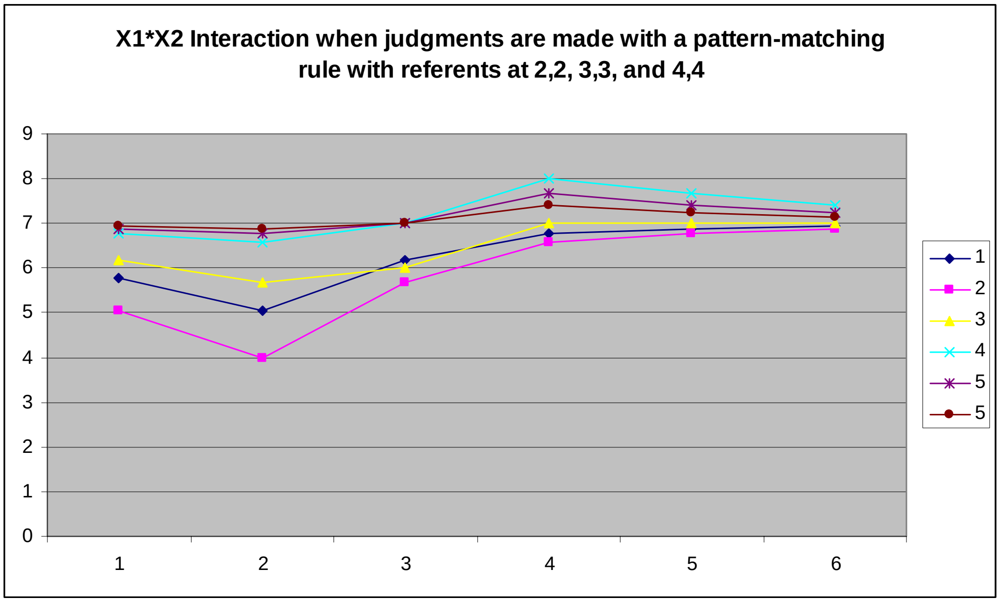


Figure 2c



Figures 2b and 2c provide good and bad news about the ability of linear models to mimic pattern-matching judgment processes. The good news is that they show as long as a consumer's previous experiences are well chosen (in this case at the extremes of utility continua), and under an appropriate (here, neutral) prior, simple linear models can do a good job of describing contemporaneous preferences *and* yield parameter estimates that are asymptotically stable[2]. Specifically, in Figure 1b we see that judgments generated by a pattern-matching process that is informed by the utility extremes will forecast the true valuations of unfamiliar intermediate profiles well, and should be well described by a linear-additive model.

The bad news is that if experiences (and/or priors) are *not* well chosen, the value of linear-models will be greatly reduced in stability and descriptive validity. As shown in Figure 1c, when referent experiences lie in the interior of the attribute space (which may be more typical in practice), the revealed preference surface becomes nonlinear,

[2] That is, recover the preferences that would be revealed under full information.

displaying an interaction at the more distant (relative to experience) tail. This implies that not only would a simple linear model do a poor job of capturing contemporaneous preferences (one needs a data design that can estimate nonlinear effects and interactions), but it also would poorly forecast the preference structure observed at future times when the consumer's scope of experience in the category expands.

This latter result holds two implications for applied choice analysts. The first is that linear-additive models will often be ill-suited for describing the association that exists between product attributes and product choices novice consumers. Differential knowledge of utility over any multi-attribute space will produce non-linearities and/or interactions that such models will fail to capture. But care must be taken in recalling that these interactions would not be manifestations enduring conditional preferences (for example, a sustained increased sensitivity to service variations given higher paid prices), but rather the transient effects of limited knowledge of preferences over the attribute space. As such, they would be expected to display little temporal stability, perhaps vanishing completely as consumers become fully knowledgeable about a product category.

The effect of unspecified variability in the unobserved components of utility.

It might be argued, of course, that because novice consumers are likely to be heterogeneous in the kinds of product experiences they have had, individual-level departures from linear-additivity due to use of naïve pattern-matching rules could well wash out when markets are viewed in the aggregate. That is, one might be able to proceed with traditional linear-additive choice models under the assumption that transient individual differences in functional form would be captured by the variance of the unobserved component of utility in a traditional linear-additive model. If the effects of subsequent learning similarly cancel themselves out across the population, then the coefficients of the linear-additive model—though perhaps wrong at the individual level—would provide good aggregate long-term forecasts of preferences.

Is the problem solved that easily? The answer, unfortunately, is “no”, for two reasons. First, because choice model estimates are perfectly confounded with the variance of the unobserved component of utility (see, e.g., Swait and Louviere 1993), changes in consumer experience that alter the structure of the unobserved component over time would also induce temporal changes in these estimates. Without a theory of what is driving the error terms, the exact nature of these temporal changes would be impossible to predict. For example, consider a choice analysis that reveals novice consumers to be statistically insensitive to variations in service quality. The confound of value and variance implies that that the meaning of this result fundamentally ambiguous: one could never know for sure whether it accrues to that fact that consumers have an enduring indifference to service, or have a sensitivity that is being temporarily masked by the aggregate effect of consumers using a heterogeneous mix of pattern-matching heuristics. If the latter is the case, parameters estimated now would be of little value for long-term planning purposes.

A second, more subtle, problem is that if a population is heterogeneous in their category knowledge, the variance of the unobserved utility component should also not be constant *across* a sample at any given point in time. Hence, inferences about preference heterogeneity derived from model parameters will be confounded with knowledge heterogeneity, or variations in the standard deviation of the unobserved component of utility across consumers.

For example, consider a case in which a population is characterized by a mixture of experienced consumers who reliably choose products based on a given set of attributes, and less experienced consumers whose choices are less reliably linked to attributes (e.g., they make judgments by referring to one of two products with which they have had direct previous experience). In both groups the consumers differ in their true sensitivity to price (that is, the sensitivity to price that would be observed if one controls for all unobserved influences on choice). So, imagine that there are four types of consumers as shown in Table 1a below: 1) low choice variability combined with low sensitivity to price and high sensitivity to quality; 2) low choice variability combined with high sensitivity to price and low sensitivity to quality; 3) high choice variability combined with low sensitivity to price and high sensitivity to quality; and 4) high choice variability combined with high sensitivity to price and low sensitivity to quality. In this table the "scale" corresponds to the inverse variance of the unobserved component of utility

Table 1a: Consumer preference & variability types

Variability X Preference	**Low price sensitivity; high quality sensitivity**	**High price sensitivity; low quality sensitivity**
Low Precision	1. $V_i = 1.5Q_i – 0.5P_i$ scale = 0.6	2. $V_i = 0.5Q_i – 1.5P_i$ Scale = 0.6
High Precision	3. $V_i = 1.5Q_i – 0.5P_i$ scale = 1.9	4. $V_i = 0.5Q_i – 1.5P_i$ scale = 1.9

Now, consider parameters estimated from these four consumer types, as shown in Figure 1b. It should be noted that "scale" multiples the systematic utility component, so the parameters in the above table are "true" parameters. Note that if we know the true parameters, we can conclude that consumer types 1 and 3 are identical and types 2 and 4 are identical, except for choice variability. Yet, with the scale confound, we would conclude that none of the four consumer types are alike, although we might incorrectly conclude that 1 and 4 share sensitivity to quality, while 2 and 3 share sensitivity to price. A similar result would obtain if instead of discrete consumer types/classes one had continuous distributions with the true parameters as means.

Figure 1b: Estimation realization from Table 1a

Variability X Preference	**Low price sensitivity; high quality sensitivity**	**High price sensitivity; low quality sensitivity**
Low Variability	1. $V_i = 0.9Q_i - 0.3P_i$	2. $V_i = 0.3Q_i - 0.9P_i$
High Variability	3. $V_i = 2.85Q_i - 0.9P_i$	4. $V_i = 0.9Q_i - 2.85P_i$

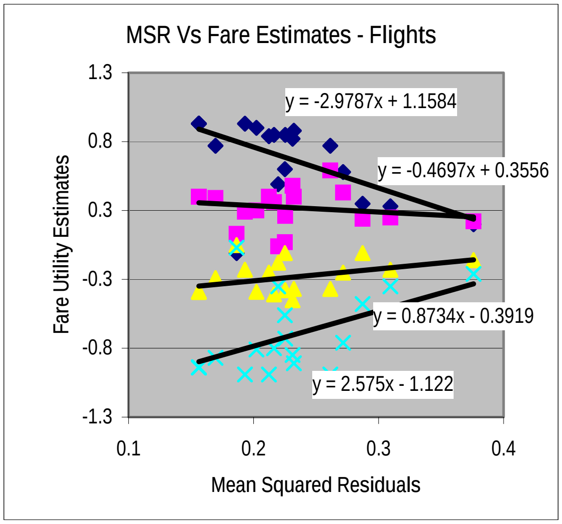
Are variance effects on model parameters real?

An obvious objection to the above discussion and stylized example is that they only establish the *possibility* that analyses of preference heterogeneity based on standard methods may be misleading. In turn, this begs the questions as to the degree to which individuals indeed differ in the variance of the unobserved utility component, and the degree to which preference parameters are confounded by this variability.

An illustration of the magnitude and nature of variance effects was recently provided by Louviere and Eagle (2006). They report the results of 66 choice experiments whereby choice models were estimated for single individuals, which allows one to estimate the size of the variance of the unobserved component of utility in choice. To provide a flavor of these analyses, in the Appendix we report the results of twenty-one individual-level model estimates from two of the 66 experiments, reflecting choices among hypothetical pizza products and cross-country flights. These two contexts are reported simply for convenience and because they have fewer parameters than other contexts. The two experiments used a common underlying optimally efficient design to estimate the main effects of a $2^3 \times 4^3$ factorial based on three options per set. Participants were members of an Australian opt-in online panel; completion rates for both conditions were over 80%.

Tables A1-a and A1-b in the appendix display the individual-level MNL estimates for the subjects who participated in the experiments, and Tables A2a and A2b give summary statistics for the experiments. The individual-level model estimates allow one to calculate residuals from model predictions, in turn allowing one to regress design matrix codes on residuals. This auxiliary regression allows one to determine if a) significant unobserved variability remains after MNL estimation, and b) the remaining unobserved variability is systematic (i.e., residuals systematically related to design elements). Both conditions produced similar results, with 18 of the 21 individuals in the pizza condition and 17 in the flights condition exhibiting regression results that are significant at the 90% C.I. Thus, the vast majority of individuals in both conditions have significant remaining unobserved variability that is systematically related to design attribute levels. Thus, it is unlikely that the individuals satisfied constant error variance assumptions.

In Figures 2a and 2b we graph the individuals' mean squared model residuals against their airfare and price utility estimates, for flights and pizzas, respectively.



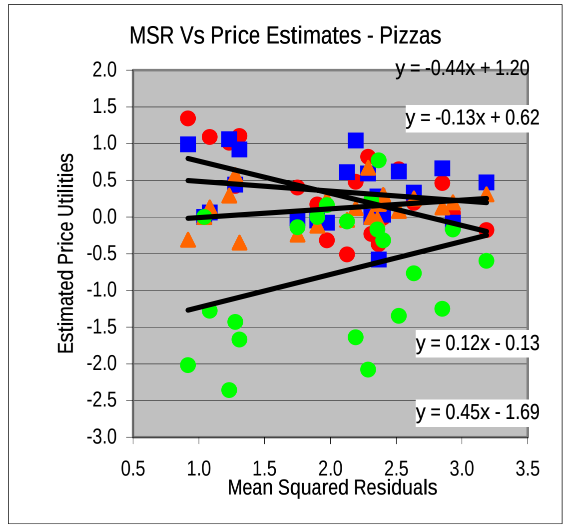


Figure 2

Both graphs are consistent with random utility theory, which predicts that as error variances increase (measured by mean square residuals), model parameter estimates should converge to zero. Both graphs display this result, allowing one to "see" that the magnitudes of the airfare and price estimates are a function of the variability in each individual's choices. Thus, in the case of airfares and pizza prices, a large proportion of parameter differences between individuals can be explained simply by differences in individuals' choice variability. Thus, model estimates of these effects are significantly confounded with individual differences in variability.

More specific details of the breakdown of these effects are as follows:

1. Summary of variance explained by choice variability (MSR) between individuals (flights)

1. Travel times - 6.8%
2. Airfares – 22.1%
3. Airline brands – 21.1%
4. Frequent flyer program – 0.4%
5. Number of stops – 24.5%
6. Free drinks – 9.1%
7. ASCs – 20.5%
8. Average across all individual estimates – 16.2%

2. Summary of variance explained by choice variability (MSR) between individuals (pizzas)

1. Pizza chain - 1.2%
2. Pizza prices – 21.2%
3. Number of toppings – 23.6%
4. Free bread – 0.0%
5. Free drinks – 8.6%
6. Free desert – 12.9%
7. ASCs – 19.3%
8. Average across all individual estimates – 13.6%

The importance of these results is that they demonstrate the danger of interpreting empirical variance in choice model parameters as uniquely reflective of either preferences or preference heterogeneity. Specifically, if true parameters are confounded with error variances, choice models can forecast future choice behavior well only if the variances are temporally stationary and/or they do not co-vary with other factors that were constant in the data source used for estimation, but are not constant at times or places or in segments to be predicted. As noted above, there is good reason to suspect that they often will not be constant, particularly in the case of developing markets, but also in many cases where no attempts are made to understand other possible sources of variability in choices (a typical case in choice modeling).

Dealing with complex choices in mature markets: short cut heuristics and their representation

The above discussion focused on problems modeling choices in early stages of market evolution when the set of market alternatives is small, but the uncertainty in how to evaluate these alternatives is high. As markets approach maturity, however, the opposite problem occurs; while there may be little uncertainty in forecasting the utility that consumers will derive from individual offerings, their choices may well be more difficult if there are a large number of differentiated offerings. In this section we take up what we know about how informed consumers make choices from complex sets, and the implications of this for standard choice analyses.

A pervasive finding of work that has studied processes associated with complex choices is that decisions are often guided by non-compensatory screening rules that act either to produce unique choice outcomes or to sequentially reduce choice sets to cognitively-manageable sizes (e.g., Payne, Bettman, and Johnson 1993). For example, individuals may eliminate alternatives if they a) fail on a critical product attribute (a conjunctive-elimination rule), b) fail to offer at least one a distinctive benefit across attributes (a disjunctive elimination rule), or c) are unattractive by virtue of their rank-order position in a set (a rank elimination rule; e.g., Einhorm 1970). In theory, the use of such decision rules could be problematic for random utility models because they imply that indirect utility functions are not strictly linear and additive. Instead, options are evaluated using non-compensatory rules that make only limited use of attribute information.

As we earlier showed, under some conditions linear models can provide a reasonably close first approximation to decisions made by pattern-matching rules. Fortunately, they also can provide a reasonably close first approximation to non-compensatory screening or elimination rules--as long as willing to consider non-additive forms that admit interactions among cues (see, e.g., Einhorn 1970). For example, consider a consumer who makes a series of binary judgments about whether each of several products described by two attributes, X_1, and X_2, is acceptable or not. The consumer makes her judgments using a non-compensatory conjunctive screening rule, as follows:

Acceptable if $[X_1>\alpha]$ *and* $[X_2>\delta]$. *(4)*

It is easy to show that a continuous algebraic analog to (4) exists as long as there is imprecision in the attribute thresholds α and δ that drive judgments in a given data set (e.g., due to momentary changes in tastes; or when calibrated from data from a sample, heterogeneity in thresholds). If α and δ act as random variables, the *probability* that a product profile (X_{i1}, X_{i2}) will be judged as acceptable is given by the product of the marginal probabilities that the realized (and unobserved) values of α and δ are less than X_{i1} and X_{i2} at the moment of choice; i.e.,

$$Pr(i|\ X_{i1},X_{i2})=Pr(Xi_1>\alpha)Pr(Xi_2>\delta). \quad (5)$$

The continuous-functional analog of (1) that follows depends on the assumed functional form of the distribution of errors associated with the acceptance thresholds α and δ. For example, if the cumulative densities of the errors associated with α and δ can be approximated by a linear probability model of the form. $\Pr(X_{i1}>\alpha)=a_0+a_1X_{i1}$ and $\Pr(X_{i1}>\delta)=b_b+b_1X_{i2}$, then data generated by a conjunctive process for generating acceptability judgments would correspond to the bilinear utility function

$$\begin{aligned} Pr(i|\ X_{i1},X_{i2}) &= (a_0+a_1X_{i1})(b_0+b_1X_{i2}) \\ &= k_0+k_1X_{i1}+k_2X_{i2}+k_3X_{i1}X_{i2}\ . \end{aligned} \quad (6)$$

(e.g., Keeney and Raiffa 1976). Expression (6) contains an important implication, namely that if the errors are associated with thresholds and the distribution of these errors is approximately linear, conjunctive screening processes are mathematically equivalent to a linear probability model that recognizes linear-by-linear interactions between attributes.

By extension, the finding of fan-like (linear-by-linear) interaction among a pair of attributes in a multi-attribute judgment experiments (e.g., a full-factorial conjoint design) has long been seen as suggesting the likely use of noisy screening rules in judgment by decision makers (e.g., Louviere 1988).

Note that we can extend this idea to any arbitrary set of conjunctive or disjunctive screening rules and error distributions. Specifically, for a set of independent acceptability judgments generated by a general family of stochastic screening rules of the form (X>,<,or=ζ), it should be clear that there will always exist an equivalent continuous algebraic counterpart of the form:

$$\Pr(i \mid X)_n = f_n(X_i) + g_n(XX_i) \qquad (7)$$

where $fn(X_i)$ is a general polynomial expansion about the attribute vector X_i, viewed by decision maker n, and $g_n(XX_i)$ is a similar expansion over the vector of cross-products or interactions among of the elements of the vector X[3].

The costs of misspecification: an empirical illustration

Expression (3) seems to provide a simple remedy for capturing choice behavior if one suspects that consumers use an unobserved array of non-compensatory screening rules. That is, simply construct an appropriate design that allows one to estimate a generalized set of interactions among product attributes. Such designs should allow analysts not only to capture the average effect of using non-compensatory rules on decisions, but also heterogeneity in the structure of these rules across a population. So, it is surprising that there has been few attempts in choice-model applications to estimate

[3] Expression (3) follows by allowing the error distribution for any screening parameter to be represented by a generalized polynomial, and by assuming that the heuristic process corresponds to a series of independent condition-act statements.

such general forms. There appear to be two reasons why this is the case. The first is pragmatic, namely that ecological data like a panel are rarely rich enough to support identification of complex models like (3). The second is that laboratory choice experiments that allow estimation impose significant data requirements that involve obtaining observations from (potentially) large factorial arrays, which historically has been seen as impractical in most field settings (although, as noted by Louviere, Hensher and Swait, 2000, this in fact is not true).

However, a more likely reason why more complex indirect utility functions are not more commonly estimated is the long-standing result by Dawes and Corrigan (1974), namely that estimating higher-order interactions adds little to model fit or out-of-sample predictive ability. Specifically, as long as attributes are monotonic in their effect on a criterion and attributes across alternatives in choice are not maximally negatively-correlated (i.e., form a perfectly efficient Pareto set), it will be the case that a strictly linear-compensatory choice model will mimick many non-compensatory choice rules (e.g., Dawes and Corrigan 1974; Einhorn, Kleinmuntz, and Kleinmuntz 1979; Johnson, Meyer, and Ghose 1988). In short, if one only cares about statistical description and prediction, simple linear models will often be good enough for who they are for.

But what if analysts are interested in more than prediction, and want to use choice models to derive insights about processes and/or the substantive nature of preferences in a population? Now, robustness no longer applies, and omitting interactions from the indirect utility function not only can lead to biased estimates, but it also can lead analysts to misleading conclusions about how product attributes influence market choices.

As an example, consider what happens if one designs a typical conjoint experiment, but instead of asking individuals to rate or rank the experimental product profiles, one asks them to evaluate each option and state if they would (yes) or would not (no) choose each. To make the example concrete, we consider pizza delivery services described by four attributes (price, brand name, number of toppings and type of crust); each attribute has two levels, and each individual is asked to evaluate and respond to the entire factorial (2^4).

We constructed 15 hypothetical individuals, each of whom is represented by a particular deterministic decision rule to say “yes” or “no” to each pizza profile. For example, an individual might use the rule “say yes if price is low and crust is thin”, or “say yes if brand is Dominos, crust is thick and number of toppings equals 4”, or “say yes if price is low and brand is Pizza Hut”. We apply the 15 rules to generate the yes’s and no’s associated with each of the 16 experimental profiles. Thus, the dataset produced by this process contains 16 scenarios x 15 individuals, or 240 observations. In the interests of space we omit typical preliminary analyses that one should conduct on the dataset, such as calculating marginal frequencies (conditional means) for each attribute level. We can summarize these analyses by noting that like almost all choice experiment datasets, the marginal frequency counts indicate that all effects are large and have acceptable signs (preference directions).

To begin our analysis of these data, we first estimate a simple, one-size-fits-all binary logit model. Model estimates and associated statistics are shown in Table 2 below. By and large all the effects are significant, although crust type is marginally significant. Instead of standard log-likelihood results, consider how well the estimated model predicts observed response probabilities. That is, each hypothetical individual faces the same 16 scenarios; hence, we can calculate the observed proportion of yes's for each scenario, and compare this with the predicted proportion of yes's from the estimated model, allowing calculation of conventional r-square values. The simple model fits the estimation data fairly well, with an r-square value of 0.73. If presented such a set of results many analysts would likely conclude that their work was done; the model fits well, and yields intuitively reasonable insights about how attributes affect choice.

Table 2

Effect	**Estimate**	**StdErr**	**Wald**	**P(wald)**
Pizzaname	-0.4911	0.1887	6.7697	0.0093
Pizzaprice	-0.5585	0.1916	8.4994	0.0036
Crusttype	0.2933	0.1832	2.5631	0.1094
Ntoppings	0.3588	0.1847	3.7758	0.0520
Constant	-1.8146	0.2067	77.0804	0.0000

But suppose we estimate an auxiliary regression on the residuals in the design matrix. Given the nature of the data generating process, it is not surprising that we obtain a highly significant regression result (F = 3.5, P(F)<0.000), with each main effect significant at the 90%C.I., and at least one interaction (price x number of toppings) also significant. So, it should be clear that something is wrong with the binary logit model.

To address this, we add all the two-way interactions to the one-size fits all binary logit, and re-estimate. This model fits the estimation data significantly better, although the price x number of toppings interaction is not significant. But there is still a problem, namely an auxilliary regression analysis of the residuals from this model again produces a significant result (F=2.5, P(F)<0.002); the main effects again are significant, and the price by number of toppings interaction again is significant. At this point, if we were presenting this finding to a conference, we would expect to hear the typical refrain of "you need to take preference heterogeneity into account". Of course, in the present case, it is not preference heterogeneity that is the source of misspecification, but *rule heterogeneity*; we have overlooked heterogeneity in the array of interactions that represent different non-compensatory rules.

As discussed above, each decision rule that was constructed can be represented as a Linear Probability Model (LPM). Each individual has a different LPM, and the LPMs generally will contain interaction terms. Table 4 below displays each of the 15 LPMs representing the decision rules. Most rules (individuals) contain one or more zero estimates, indicating that these particular effects are not part of the rule. The table is divided into two parts: a) the left-hand side contains estimated main effects, with r-squares for each individual in column seven; and b) the right-hand side contains all two-way interactions, with associated r-squares for all main effects and two-way interactions in the last column. It should be obvious that all r-squares increase substantially when we add interactions. Individuals with an r-square value of 1.0 are fully described by a rule that requires only main effects and two-way interactions; individuals with r-square values less than one require additional interactions that we omit to save space.

Table 4: LPM Estimates for 15 Rules													
					Num of	Main Effects	brand	brand	brand	price	price	crust	Main Eff+Int
Ind	Const	Brand	Price	Crust	Tops	R-Sq	price	crust	ntops	crust	ntops	ntops	R-Sq
1	0.13	-0.13	-0.13	-0.13	0.00	0.43	0.13	0.13	0.00	0.13	0.00	0.00	0.86
2	0.25	-0.25	0.00	0.00	-0.25	0.67	0.00	0.00	0.25	0.00	0.00	0.00	1.00
3	0.25	-0.25	-0.25	0.00	0.00	0.67	0.25	0.00	0.00	0.00	0.00	0.00	1.00
4	0.06	0.06	-0.06	0.06	0.06	0.27	-0.06	0.06	0.06	-0.06	-0.06	0.06	0.67
5	0.06	-0.06	-0.06	0.06	0.06	0.27	0.06	-0.06	-0.06	-0.06	-0.06	0.06	0.67
6	0.13	0.00	-0.13	-0.13	0.13	0.43	0.00	0.00	0.00	0.13	-0.13	-0.13	0.86
7	0.25	0.25	0.00	0.25	0.00	0.67	0.00	0.25	0.00	0.00	0.00	0.00	1.00
8	0.13	-0.13	0.00	-0.13	0.13	0.43	0.00	0.13	-0.13	0.00	0.00	-0.13	0.86
9	0.25	-0.25	0.00	0.25	0.00	0.67	0.00	-0.25	0.00	0.00	0.00	0.00	1.00
10	0.13	0.00	-0.13	0.13	0.13	0.43	0.00	0.00	0.00	-0.13	-0.13	0.13	0.86
11	0.25	0.00	-0.25	0.00	0.25	0.67	0.00	0.00	0.00	0.00	-0.25	0.00	1.00
12	0.25	0.00	-0.13	0.00	0.00	0.08	-0.13	0.00	0.25	-0.13	-0.13	0.00	0.67
13	0.25	0.00	0.25	0.25	0.00	0.67	0.00	0.00	0.00	0.25	0.00	0.00	1.00
14	0.13	-0.13	-0.13	0.00	0.13	0.43	0.13	0.00	-0.13	0.00	-0.13	0.00	0.86
15	0.06	-0.06	-0.06	-0.06	0.06	0.27	0.06	0.06	-0.06	0.06	-0.06	-0.06	0.67

This analysis gives us a much richer and more accurate view of what drives choices in this market. For example, it highlights that various product attributes affect choice not as independent main effects but rather as interactions with other attributes, and also that the pattern of these interactions varies considerably within the population.

Other approaches to capturing rule heterogeneity.

It is important to emphasize that this is but one of a number of approaches that have been suggested over the years for representing individual non-compensatory choice process in choice models, each having its own comparative strengths and weaknesses (e.g., Gilbride and Allenby 2004). For example, a major downside of representing non-compensatory heuristics using generalized families of interactions as illustrated above is that interaction parameters do not have a direct translation to specific noncompensatory heuristics. Moreover, interpretability of parameter estimates is further confounded by the existence of other more mundane sources of interaction effects on choices, such as levels of quality being interpreted differently depending on the observed price.

Alternative proposals for capturing rule heterogeneity that could avoid these ambiguities have been provided by (*inter alia*) Elrod, Johnson, and White (2005), Gilbride and Allenby (2004), and Swait and Adamowicz (2001), who describe generalized choice models that recognize the existence of a mix of compensatory and non-compensatory choice heuristics. However, a limitation of these proposed approaches is that they capture variation in only a small set of pre-specified heuristics (e.g., compensatory versus conjunctive screening rules). If consumers make decisions using a mix of rules other than those assumed by these models (likely to be the case), their value as descriptive (and possibly predictive) tools would likely be less obvious.

The problem of consumer foresight

The final topic we take up is a challenge to choice modelers in all phases of market evolution, namely that consumers make choices with knowledge of and expectations about future consequences. An often-cited limitation of standard choice analyses is that they assume consumers are not only utility maximizers, but *myopic* utility maximizers whereby their goal is to choose that option that offers the highest expected pleasure at the time of choice, without thinking about how this choice may affect the utility/pleasure associated with future choices (e.g., Erdem and Keene, 1996). Thus, few standard analyses allow for the possibility that a consumer may choose an option for the mere purpose of gathering information about it, delay choice out of a belief that a better choice set will be available at a later point in time, or elect not to choose an obviously superior option in order to savor the anticipation of its later consumption. Naturally, standard models have no problem *describing* choices in such settings; a decision to delay consumption, for example, can be well described by a model that posits low utility for the good at the time of choice. But such a mere statistical description clearly is dissatisfying because it ignores dynamics that produce the behavior (the distinction between not choosing and delaying choosing); and such a model only can predict behavior well in contexts identical to that in which it was estimated.

A perhaps more serious consequence of overlooking dynamics is that reduced-form or cross-sectional models of dynamic processes often will yield parameter estimates that suggest that consumer decisions are less rational than they really are. For example, a classic finding involves upward sloping contemporaneous demand curves. If consumers believe that prices set by sellers will be higher tomorrow than they are today, short-term price increases may display positive elasticities (see, e.g., Erdem, Imai, and Keane 2003). The reason is not that consumers prefer higher prices to lower prices, but instead their buying decisions are made in light of their beliefs about what *future* prices will be, which can give the appearance of a positive short-term reaction to observed price increases.

Due to increased recognition of these concerns, a major growth area in choice modeling research in marketing has been dynamic structural models that explicitly assume that consumers are multi- rather than single-period choice optimizers (e.g., Erdem and Keene 1996; Gonul and Srinivasan 1996; Erdem, Imai, and Keane 2003; Song and Chintagunta 2003). General acceptance of this work, however, has been limited by two factors. One factor is the pragmatic problem of computational complexity, such that efficient ways to empirically solve complex dynamic optimization problems have become available only recently, and are not part of the standard set of estimation tools familiar or available to applied choice modelers. This limitation may only be temporary, but the second factor is more basic, namely, the complexity of such models is viewed by some as an unreasonable behavioral description of consumer planning (Houser, Keane, and McCabe 2004). Specifically, given the well-known finding that individuals find it difficult to make once-off decisions in an optimal manner (e.g., choosing gambles in the way prescribed by expected utility theory), intuition suggests that they would have little chance to optimally solve more complex dynamic programming problems. Yet, this is exactly what dynamic structural models assume that consumers are able to do.

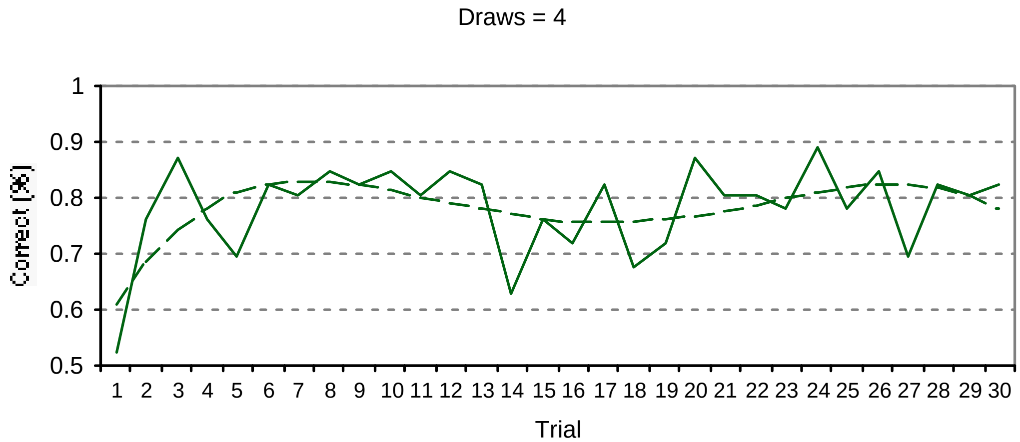
But is this intuition correct? It is important to remember that in economics the acid test of whether a given model is theoretically tenable is *not* whether it is cognitively realistic (probably few are), but instead whether it describes equilibrium behavior that could be reached by evolutionary processes. That is, boundedly-rational decision makers need only be more prone to repeat actions that tend to give higher payoffs. The fact that consumers make little attempt to plan ahead or have no idea what “backward induction” means does not preclude them from acting *as if* they do.

As an example of this, Hutchinson and Meyer (2005) recently reported the results of a study examining the ability of people to make accurate judgments about the expected maxima of distributions, an ability assumed in most optimal dynamic decision models (e.g., job search models). They studied this in two related experiments. The first was a paper-and-pencil task in which participants asked to provide intuitive estimates of the expected value of the largest number that would be realized from *N* draws form a 0-100 uniform distribution, where *N* varied from 2 to 10. Perhaps not surprisingly, participants did not perform this task well, displaying a consistent tendency to underestimate the true maxima. For example, for the case of two draws (the simplest case), where the normative answer is 67, the mean estimate was 58 with individual guesses ranging from 50 to 65. Given this result, one might think that it reasonable to question any model that assumes that people are good intuitive judges of maximum order statistics. But would that be correct?

To answer this, a second experiment was conducted in which participants were asked to play a lottery game for money that required implicit rather than explicit knowledge of expected maxima. Specifically, subjects played a computer game that required them to place a bet on which of two sets of *N* draws from a uniform urn would yield the higher maximum. In the game participants first observed a simulated dealer take *N* draws from the urn (*N* varied across tasks). After seeing the maximum value drawn by the dealer, they were then asked to indicate whether this value was likely to be higher than that which would be realized given *N* new draws. After participants made this prediction, *N* new draws were taken, and the outcome revealed. If their directional bet proved correct they received a small cash payment, but they received nothing if their bet proved wrong. Participants placed 30 such bets within each of three draw-size conditions (2, 4, and 6).

While participants may not be able to compute the maxima of distributions, it turns out that they can play *as if* they can. Across all three N-size conditions and trials, participants placed the normatively-correct bet more than 80% of the time. In Figure 3, taken from Hutchinson and Meyer (2005), we graph the proportion of correct guesses over trials for the *N=4* condition (the others look similar):

Figure 4



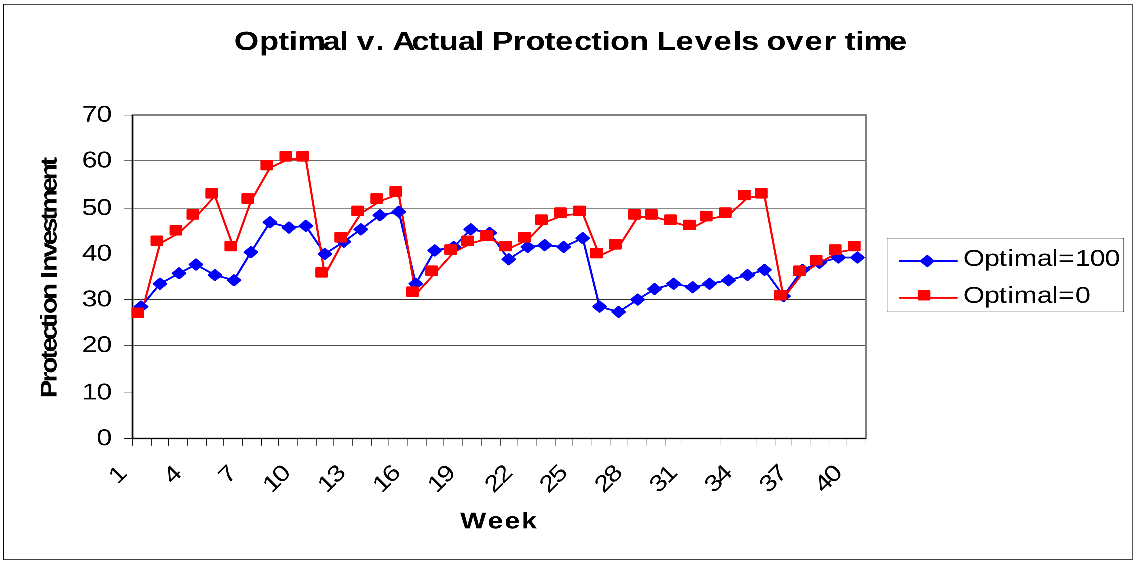
What is surprising about these data is that they not only reveal that subjects had a good implicit knowledge of maxima, but that they acquired this "as-if" knowledge quite rapidly. In many cases a single experience with feedback seemed to be enough to do the trick (Figure 4). Hence, much like the famous example of pool players and physics, the requirement for behavior to be optimal is not that people can compute optima, but rather that they live in a world that naturally reinforces optima.

But such optimism about "as if" optima comes with a strong word of caution, namely that the world does not provide consumers with frequent opportunities to learn, or reliably rewards optimal behavior. For example, in many cases the optimality of strategic choice policies cannot be observed in the short run, but only by their repeated application over a long time horizon (such as dieting). The more short-sighted a consumer, or the more stochastic the short-term reward, the more she is likely to stray from optimality by basing decisions on tangible short-term consequences rather than less-tangible long-term ones.

To illustrate this, Meyer and Kunreuther (2005) describe the results of a computer simulation designed to assess the ability of people to make optimal decisions to invest in long-term protection against a low-probability, high-consequence hazard, which in this case is an earthquake. In the simulation participants were asked to imagine that they would be living for five years in a country that was prone to periodic earthquakes. They were given a home with a certain fixed value, and as time elapsed in the simulation they had the option to invest some of this home value in permanent home improvements that would potentially lower the home's vulnerability to earthquake damage or instead to invest it in a bank at a fixed interest rate. At the end of the simulation they were paid an amount tied to original home value minus losses due to earthquake damage and investments in protection.

The central manipulation was whether or not these investments in protection were truly effective. Participants were told that there was a 50% chance that the investments did little to reduce damage, and a 50% chance that they would be highly effective. To allow learning, they played eight rounds of the simulation during which they had an opportunity to observe not only the damage that their own home suffered from earthquakes conditional on their investments, but also the investments and damage made associated with other players.

In Figure 5 we graph the actual percentage investments over each five-year block of the simulation by the optimal level of mitigation (conditional on knowing its true effectiveness).



The figure shows a disturbing result, namely that not only were the participants unable to learn the investment optimum, but the mean investment level was systematically *higher* when investments did nothing to lower earthquake losses.

What explains this dysfunctional pattern of behavior? The answer is twofold. First, a statistical analysis of the period-by-period investment decisions indicated that the primary driver of decisions to invest was the magnitude of loss actually experienced in the previous period. Hence, effective mitigation had a paradoxically *negative* effect on participant's perceived urgency to invest more when it was really effective. The second reason is that because quakes were infrequent, the benefits of protection were rarely immediately seen. Instead, it was more common for participants who invested in protection to find out that it was not immediately needed than to find out it was immediately useful, a factor that further suppressed interests in buying.

The contrasting nature of these two examples underscores the fact that the choices we observe in markets are much more a reflection of the structure of the environmental feedback received by consumers than intuitive math abilities. The optimistic implication of this for applications of dynamic structural models that criticism of their validity as a literal account of how decisions are made is largely irrelevant to their status as useful empirical theories. More precisely, the fact that one the component mathematical assumptions of a model (e.g., that people can accurately judge order statistics) is empirically invalid does not imply empirical failure of the holistic predictions of an associated model. Rather, that outcome depends on something different, namely whether consumers can observe feedback from the choices they make in markets, and whether the nature of this feedback in the long run will favor economically optimal behaviors over suboptimal ones.

Discussion

This paper was motivated by a desire to enhance cross-learning among the three traditionally disparate behavioral, economic, and statistical approaches to studying individual choice behavior. Central to the discussion was the suggestion that at best most current choice models offer a static snapshot of preferences in markets. The estimated parameters of choice models reflect not just the enduring preferences that consumers may have for products and attributes, but also the momentary state of consumer learning about products and/or heuristic short cuts they use to deal with the complexity of markets. That is, the models provide something akin to a one-dimensional view of a multi-dimensional process. In our view the key to building better planning models is not to build ever-more complex and accurate statistical descriptions of such a one-dimensional projection, but instead we should be trying to better understand the multi-dimensional process that underlies the projection.

Although this paper focused on how behavioral theory can help choice modelers gain better understanding, it is important to emphasize that a reverse flow can be no less valuable. That is, economic theory may be useful in allowing behavioral researchers to build better models for explaining the simple mechanics that actually underlies choice anomalies or laboratory findings where behavior appears to depart from the prescriptions of rational choice theory.

Reversing the dialogue: an example

To illustrate, consider multi-attribute choice models developed to account for an anomaly in choice behavior known as the *compromise effect* (Tversky and Simonson 1993). A simplified account of the effect is that when participants in choice experiments are shown an array of options arrayed along a Pareto frontier in a multi-attribute space, there is a tendency for the aggregate choice portions to be massed toward the center, regardless of where the choice options are aligned along the frontier. Hence, one can increase the odds of an option being chosen simply by framing it as the compromise alternative in a set.

This effect is termed an anomaly because, taken at face value, it violates the fundamental property of random utility theory known as regularity. That is, one should not be able to increase the odds of choosing an alternative in a choice set by enlarging the set (i.e., the choice probabilities should obey regularity). Yet, studies of the compromise effect suggest that just such an effect is possible, namely that one can increase the choice share of an extreme option by introducing a new option that is even more extreme.

Several researchers have observed that this effect can be reconciled within a random-utility framework by assuming that consumers make choices using strict utility functions where attribute values are assessed vis-à-vis choice set-specific extremes (e.g., Tversky and Simonson 1993; Kivetz, Oded, and Srinivasan, 2006; Shen, Parker, and Nakamoto 2005). For example, Kivetz, Oded, and Srinivasan, (2004) showed that compromise-effect data are well fit by a "contextual concavity model":

$$U_{nu} = b_i + \sum_k b_k (x_{ik} - x^S{}_{\min,k})^{c_k} \qquad (8)$$

Where $x^S{}_{\min,k}$ is the smallest observed value of attribute k within choice set S and c_k is an empirical shape or concavity parameter for attribute k (Kivetz , Oded, and Srinivasan 2004; Shen, Parker, and Nakamoto 2005 offer similar forms).

Does expression (8) provide a useful theoretical account of the compromise effect? While there is ample evidence that it offers a good *statistical description* of the effect, its value as a *theoretical explanation* is less obvious. The central issues in that the compromise effect not a universal phenomenon, but rather is observed only under restricted laboratory conditions where:

1. Participants are uncertain how to value and trade off attributes (the effect does not work, for example, for choices among mixtures of monetary payoffs); and
2. Choices are made by different groups of subjects viewing different choice sets with no feedback.

Preference uncertainty, however, is not explicitly modeled in (8), making it an incomplete account of the phenomenon. While the model can statistically describe compromise effects in laboratory tasks designed to create it, it cannot endogenously predict what would happen if we were to alter some of the basic conditions of the task, such as reducing uncertainty through learning.

How might one build a model of the task where uncertainty is endogenous? One possible—though unlikely--starting point is to assume that participants deal with preference uncertainty in the way that would be prescribed by rational theories of risky choice. It turns out that doing so leads to a surprising result: not only we are led to a model that endogenously recognizes preference uncertainty, but also reveals that the compromise effect may not be an "anomaly" at all.

To see this, imagine that you are invited to play a gamble in which you are offered three options, each described by a value on two attributes that are expressed in arbitrary units of measurement called "*ps*" units and "*kz*" units:

	Option		
	A	B	C
Attribute 1	75ps	50ps	25ps
Attribute 2	2kz	5kz	8kz

Each of the units of measurement has a linear rate of conversion to a dollar payoff, but the nature of this conversion is unknown. Specifically, you only know that for each attribute *i* there is a payoff $P_i=a_i+b_iX_i$ where P_i is the payoff in dollars, X_i is the observed value of *i* (expressed in units of *ps* or *kz)*, and a_i and b_i are realizations of random variables with joint distribution $f(a_i,b_i)$, such that $b_i \geq 0$. Your goal is choose the option that delivers the highest joint payoff across both attributes.

Although highly abstract, the task should be recognized as capturing the essential uncertainty that participants face in compromise experiments. One is asked to make a choice among options in which one is unsure about the mapping that exists between attribute values and utility (other than more is better), and where one has no opportunity to learn these tradeoffs by choice experience.

Is there a rational solution to this problem? There is, and is actually quite simple. First, because the original units of measurement are arbitrary, and constant scale differences *between* attributes do not affect the solution to the choice problem, we can reduce the dimensionality of the joint distributions $f(a_i, b_i)$ by rewriting the matrix above in an equivalent normalized form; i.e.,

	Option		
	A	B	C
Rank (Attribute 1)	1	.5	0
Rank (Attribute 2)	0	.5	1

If $g(b_i)$ is the resulting marginal distribution of b_i, each option thus has expected payoff

$$E(P_i) = v(Z_{i1}\int_{b_1} b_1 g(b_1)db_1) + v(Z_{i2}\int_{b_2} b_2 g(b_2)db_2 \qquad (9)$$

$$= v(Z_{i1}\overline{b_1}) + v(Z_{i2}\overline{b_2}) \qquad (10)$$

Where Z_{ij} is the normalized score of option i on attribute j, and $v(\,)$ is the decision maker's marginal value function over money.

Expression (9) makes a simple (and quite intuitive) prediction: for the current example were $Z_{i1+} Z_{i2}=1$, under neutral priors (i.e., $\overline{b_1} = \overline{b_2}$), a risk-neutral decision maker for whom $v(\,)$ is linear would be indifferent among the options. That is, he or she would recognize that there is no one best answer to the problem as long as the attribute-payoff conversions are unknown. On the other hand, if the decision maker were *risk* *averse*; that is, the person has a value function that is strictly concave over $Z\bar{b}$, then she should pick the middle or compromise option[4].

Now, here is the critical step. In the typical experimental set-up associated with demonstrations of the compromise effect (see, e.g., Kivetz, Oded, and Srinivasan 2004) an experimenter presents *two different groups of* participants *different but overlapping* choice sets in which the option that was previously the compromise is now displayed as either a high or low extreme. For example, imagine in our case instead of the gamble above you were *initially* shown the set

	Option		
	A	B	C
Attribute 1	50ps	25ps	0ps
Attribute 2	5kz	8kz	10kz

If you had indeed initially been shown this set, which option would be the rational choice? The answer is not the 50/5 option that was the compromise in the last set, but, if one is risk averse, the 25/8 option that is the compromise alternative in this new set. The reason is this: *Because the attribute scales have no absolute metric in a payoff (or utility) space*, and since no learning is possible, the normative analysis is exactly the same for both the old and new choice sets. An assumption that individuals are risk averse over uncertain preferences leads to a predicted bias toward choosing the middle option on a Pareto frontier in both cases.

Hence, what the compromise effect shows is *not* that people do not make choices in a way that is consistent with utility theory, but rather that *one assumption* that

[4] This follows from the definition of concavity, which requires [v(.5k)+v(.5k)]>[v(0k)+v(1k)]

accompanies typical applied analyses may not always hold. That is, the assumption that attribute levels have a constant absolute meaning (and mapping to utility) across a range of choice sets. If one relaxes that assumption and builds a model that formally recognizes uncertainty in utility exchange rates, compromise data can be easily reconciled with standard theory. The effect *seems* anomalous only because readers have access to holistic knowledge about the range of attributes that participants in the experiments did not.

Are the models suggested by Kivetz, *et al* wrong? To the contrary, the above analysis ironically leads us to the same conclusion we reached about how one can algebraically *describe* compromise effects. That is, their contextual concavity model (expression (8)) can be directly motivated as a model of risk averse preferences for consumers who are unsure about the scaling of attributes and subjectively normalize them over the range displayed in the experiment. The critical difference is that by explaining the result in terms of its origins in risky decision making we can endogenously predict the model's failure. Expression (10) implies that preferences for compromise options should vanish as uncertainty in preferences (*var(b))* diminishes and/or individuals are exposed to choice sets with broader ranges sequentially.

Conclusion

It would be wrong to conclude that the above illustration implies that rational models of choice enjoy any higher status as tools of explanation than any other type of model (behavioral or statistical), or that behavioral researchers should reject all possible normative benchmarks (however far-fetched) before deriving their own explanation. In this case a rational model of risky choice was appropriate for entirely pragmatic reasons; it offered a convenient representation that satisfied the theoretical modeling requirements of simplicity and endogenous recognition of uncertainty. Were we to continue this modeling effort, ideally the flow of dialogue would again reverse, with behavioral research being asked to refine the risky choice model to better reflect the realities of how consumers actually deal with preference uncertainty and to incorporate lay beliefs about the scaling and benefits of attributes (e.g., Machina 1982).

As we mentioned at the outset of the paper, our goal was to foster more dialogue among what have become increasingly disparate approaches to understanding and modeling decisions and choices. In recent years behavioral research has made significant advances in providing better understanding of how consumers learn preferences and make choices, but without a clear connection to empirical choice modeling, either in terms of how efforts can be improved by this knowledge or what the consequences are of ignoring it. In this paper we tried to illustrate in a limited way how to begin to build such bridges. We hope it represents only the first words in what should prove to be a long dialogue.

References

Abe, Makoto (1995), “A Nonparametric Density Estimation Method for Brand Choice Using Scanner Data”, *Marketing Science* 14,(3); 300-325.

Allenby, G.M. and P.E. Rossi (1999), "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics*, 89 (1-2), 57-78.

Bettman, J.R., E.J. Johnson, and J.W. Payne (1991), "Consumer Decision Making," in *Handbook of Consumer Behavior*, T.S. Robertson and H.H. Kassarjian (eds.), Englewood Cliffs, NJ: Prentice-Hall.

Camerer, Colin, and Teck Ho, (1999*), "*Experience-Weighted Attraction Learning in Normal Form Games,*" Econometrica,* 67 (1999), 837-874.

Dawes, Robin, and Berbard Corrigan (1974), “Linear Models in Decision Making”, *Psychological Bulletin,* 81, 05-106.

Elrod, T., R.D. Johnson and J. White (2005), “A new integrated model of noncompensatory and compensatory decision strategies,” Organizational Behavior and Human Decision Processes, 95 (1), 1-19.

Einhorn, H.J. (1970), "The Use of Nonlinear, Concompensatory Models in Decision Making," *Organizational Behavior and Human Performance*, 73, 221-230.

Einhorn, H. J., D. N. Kleinmuntz, D. N., and B. Kleinmuntz, B. (1979). Linear regression *and* process-tracing models of judgment. *Psychological Review*, 86, 465–485.

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

Erdem, Tülin, Imai, Susumu, and Michael P. Keane (2003), “A Model of Consumer Brand and Quantity Choice Dynamics under Price Uncertainty,", *Quantitative Marketing and Economics* 1 (1), 5-64.

Fudenberg, Drew, and David K. Levine (1998), *The Theory of Learning in Games.* Cambridge, MA: the MIT Press.

Gilboa, Itzhak, and David Schmeidler(1995), "Case-Based Decision Theory", *Quarterly Journal of Economics*, November, 605-630.

Gilbride, T.J. and G.M. Allenby (2003), "A Choice Model with Conjunctive, Disjunctive,and Compensatory Screening," *Marketing Science*, forthcoming.

Gonul, Fusun, and Kannan Srinivasan (1996), “Estimating the Impact of Consumer Expectations od Coupons on Purchase Behavior: A Dynamic Structural Model”, *Marketing Science*, 15(3), 282-280.

Jennifer Gregan-Paxton, Deborah Roedder John (1997), “Consumer learning by analogy: A model of internal knowledge transfer”, *Journal of Consumer Research*, 24 (3), 266-287.

Hensher, David A., and William H. Greene (2003) "The Mixed Logit Model: the State of Practice", *Transportation*, 30, 133-176.

Houser, Daniel, Keane, Michael, and Kevin McCabe (2004) “Behavior in a dynamic decision problem: An analysis of experimental evidence using a Bayesian type classification algorithm,”. *Econometrica*, 72:3 (May), 781-822.

Hutchinson, J. Wesley, and Robert J. Meyer (2005), “Extreme Bias: Heuristics and Biases in Intuitive Judgments about Maximum-order Statistics”, *working paper*, Department of Marketing, the Wharton School, University of Pennsylvania.

Johnson, E.J., R.J. Meyer, and S. Ghose (1989), "When Choice Models Fail: Compensatory Models in Negatively Correlated Environments," *Journal of Marketing Research*, **26** (3), 255-270.

Kamakura, W.A. and M. Wedel (2004) “An Empirical Bayes Procedure for Improving Individual Level Estimates and Predictions from Finite Mixtures of Multinomial Logit Models,” *Journal of Business and Economic Statistics*, 22 (1), 121-126.

Kamakura, W.A., B.D. Kim, and J. Lee (1996), "Modeling Preference and Structural Heterogeneity in Consumer Choice," *Marketing Science*, **15** (2), 152-172.

Keeney, R.L. and H. Raiffa (1976), *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, New York: Wiley.

Meyer, Robert J. and Howard Kunreuther (2005), “Earthquakes in the Lab: the Optimality of Decisions to Invest in Long-Term Mitigation”, *working paper*, Center for Risk and Decision Processes, University of Pennsylvania.

Louviere, J.J. (1988), *Analyzing Decision Making: Metric Conjoint Analysis*, New York: Sage.

Louviere, J.J. and T.C Eagle (2006) “Confound it! That Pesky Little Scale Constant Messes Up Our Convenient Assumptions!” Proceedings, Annual Sawtooth Software Conference, Delray Beach Florida, March (http://www.sawtoothsoftware.com/techpap.shtml).

Louviere. J., D. Hensher and J. Swait (2000) *Stated Choice Methods: Analysis and Applications.* London: Cambridge University Press.

Machina, Mark (1982), “Expected Utility Analysis without the Independence Axiom,” *Econometrica* 50 (March 1982) 277-323.

Manchanda, Puneet, Rossi, Peter E., and Pradeep Chintagunta (2004), “Respose Modeling with Nonrandom Marketing Mix Variables”, Journal of Marketing Research, 41(4), 467

Mellers, Barbara E. (1981), “Configuarlity in Multiple-Cue Learning”, *American Journal of Psychology*, 93, 429-443,

McFadden, Daniel (1981), "Econometric Models of Probabilistic Choice," in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Duscrete Data with Econometric Applications*, , MIT Press: Cambridge, MA, 198-272

Norman, Donald A. (1988), *The Psychology of Everyday Things,* New York: Basic Books.

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

Rossi, Peter E., and Greg M. Allenby (2003), “Bayesian Statistics and Marketing”, *Marketing Science*, 22(3), 304-328.

Rossi, P.E., Allenby, G.M. and R. McCulloch (2005) *Bayesian Statistics and Marketing*, John Wiley and Sons.

Sheng, Shibin, Parker, Andrew M., and Kent Nakamoto (2005), Understanding the Mechanism and Determinants of Compromise Effects**”,** *Psychology and Marketing*, 22(7), 591-610.

Song, Inseong and Pradeep Chintagunta (2003), “A Micromodel of New Product Adoption with Heterogeneous and Forward-Looking Consumers: Application to the Digital Camera Category”, *Quantitative Marketing and Economics*, 1(4), 371.

Swait, Joffre, and Wictor Adamowicz (2001), "The Effect of Task Complexity on Consumer Choice: a Latent-Class Model of Decision Strategy Switching", *Journal of Consumer Research,* 28 (June), 135-148.

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

ter Hofstede, F., Kim, Y. and M. Wedel (2002) “Bayesian prediction in hybrid conjoint analysis,” *Journal of Marketing Research*, 34 (2), 253-261.

Tversky, A. (1972), "Elimination by Aspects - Theory of Choice," *Psychological Review*, **79** (4), 281-299.

Tversky, A. and I. Simonson (1993), "Context-Dependent Preferences," *Management Science*, **39** (10), 1179-1189.

Wedel, Michel and Wagner A. Kamakura (1999), *Market Segmentation: Conceptual and Methodological Foundations*, 2nd edition, Boston: Kluwer Academic Publishers.

Appendix: Individual Logit Estimation Results

Table A1a: Individual-Level MNL Model Estimates for Flights																		
ID	MSR	4 hrs	5 hrs	6 hrs	7 hrs	$350	$450	$550	$650	Qant	Vblue	Jstar	Aust	FFlyer	Stops	Drinks	asc1	asc2
1	0.662	-0.060	-0.150	0.200	-0.452	-0.150	0.220	0.000	-0.070	1.870	0.760	-0.900	-1.730	0.090	-0.050	0.050	1.630	1.350
2	0.864	0.080	0.050	0.040	-0.994	1.770	0.860	-0.410	-2.220	0.150	-0.230	0.000	0.080	0.220	0.240	-0.230	1.370	1.520
3	0.791	0.410	-0.030	-0.130	-1.171	1.900	0.690	-0.340	-2.250	-0.310	0.220	-0.010	0.100	0.010	-0.030	0.130	1.480	1.350
4	0.534	0.000	0.000	0.000	-0.534	1.880	0.860	-0.710	-2.030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.600	1.300
5	0.534	0.000	0.000	0.000	-0.534	1.880	0.860	-0.710	-2.030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.600	1.300
6	0.806	0.290	-0.030	-0.290	-1.066	1.730	0.710	-0.760	-1.680	-0.110	0.100	-0.050	0.060	-0.100	0.120	-0.010	1.750	1.510
7	1.014	0.760	0.020	-0.070	-1.794	1.820	0.830	-0.680	-1.970	-0.470	0.290	0.380	-0.200	-0.030	-0.030	0.090	1.550	1.300
8	1.235	0.800	0.960	-0.180	-2.995	1.170	1.020	-0.530	-1.660	-0.120	0.120	-0.010	0.010	0.130	-0.020	-0.280	1.200	1.590
9	1.692	-0.070	0.330	0.380	-1.952	0.640	0.470	-0.420	-0.690	0.250	1.400	-2.720	1.070	0.070	0.040	0.040	1.110	1.350
10	0.534	0.000	0.000	0.000	-0.534	1.880	0.860	-0.710	-2.030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.600	1.300
11	1.310	0.640	0.010	-0.510	-1.960	1.610	1.310	-0.570	-2.350	0.390	-0.080	0.010	-0.320	-0.080	0.080	-0.340	1.410	1.510
12	0.924	-0.200	-0.100	-0.110	-0.624	1.770	0.730	-0.850	-1.650	-0.180	-0.200	0.000	0.380	0.180	-0.150	-0.090	1.830	1.400
13	0.534	0.000	0.000	0.000	-0.534	1.880	0.860	-0.710	-2.030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.600	1.300
14	1.520	0.700	0.330	-0.800	-2.550	0.590	0.580	-0.030	-1.140	-0.160	0.020	-0.040	0.180	-0.110	0.570	0.010	2.320	1.720
15	0.534	0.000	0.000	0.000	-0.534	1.880	0.860	-0.710	-2.030	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.600	1.300
16	1.022	1.000	0.070	-0.160	-2.092	1.670	1.020	-0.840	-1.850	-0.460	0.280	0.180	0.000	0.030	0.090	-0.040	1.360	1.470
17	1.057	0.500	0.230	-0.430	-1.787	1.690	0.670	-0.860	-1.500	-0.010	0.270	-0.010	-0.250	0.000	0.230	-0.200	1.690	1.790
18	1.038	0.630	-0.190	-0.210	-1.478	1.090	0.280	-0.070	-1.300	0.200	0.190	-0.370	-0.020	-0.080	0.290	-0.680	1.830	1.660
19	0.911	1.410	0.920	-0.490	-3.241	1.100	-0.050	-0.320	-0.730	-0.230	0.000	0.150	0.080	0.000	-0.240	-0.080	1.570	1.480
20	2.197	0.200	0.050	0.050	-2.447	0.160	0.600	-0.220	-0.540	0.510	0.990	-1.940	0.440	0.010	0.610	-0.530	1.430	2.000
21	0.614	-0.080	0.300	-0.120	-0.834	1.500	0.810	-0.500	-1.810	0.270	0.060	-0.090	-0.240	-0.140	-0.300	-0.350	1.520	1.690
Abbreviations: MSR=Mean squared residual; Qant=Qantas; Vblue=Virgin Blue; Jstar=JetStar; Aust=Australian Airlines; Fflyer= Frequent Flyer program; Stops=Number of Stops; Drinks=Free wine/beer; asc = alt-specific constant; hrs = flying time; $ = fare																		

Table A1b: Individual-Level MNL Model Estimates for Pizzas

ID	MSR	Phut	Dom	Boys	Phav	$12	$14	$16	$18	ntop1	ntop2	ntop3	ntop4	bread	drink	desrt	asc1	asc2
1	2.287	0.700	0.200	-0.480	-3.187	0.820	0.590	0.670	-2.080	-1.310	0.260	0.770	0.280	-1.020	-0.130	-0.220	1.280	1.060
2	2.519	0.280	-0.380	0.010	-2.419	0.650	0.620	0.080	-1.350	-1.510	0.560	0.070	0.880	-0.530	-0.570	-0.290	1.790	1.740
3	1.279	-0.740	0.180	0.010	-0.719	0.420	0.440	0.570	-1.430	-1.630	-0.170	0.840	0.960	-0.230	-0.260	0.000	1.710	1.570
4	2.127	-0.610	0.010	1.120	-1.527	-0.510	0.610	-0.040	-0.060	-1.980	0.070	1.350	0.560	0.050	-0.260	-0.430	1.560	1.530
5	1.750	-0.300	0.300	0.000	-1.750	0.400	-0.020	-0.240	-0.140	-1.680	-0.480	0.830	1.330	-0.220	-0.660	0.180	1.840	1.160
6	2.932	0.100	-0.080	-0.710	-2.952	0.050	-0.080	0.200	-0.170	-3.020	-0.100	1.510	1.610	0.110	-0.260	0.030	1.360	1.040
7	3.186	-0.770	0.730	0.290	-3.146	-0.180	0.470	0.310	-0.600	-2.430	0.190	1.120	1.120	-0.430	-0.800	-0.520	1.590	1.240
8	2.310	1.210	2.310	-3.060	-5.830	-0.230	0.000	0.000	0.230	0.090	0.210	-0.230	-0.070	-0.150	0.150	-0.100	0.880	0.920
9	1.310	-0.070	-0.210	0.170	-1.030	1.100	0.920	-0.350	-1.670	-1.040	1.030	0.210	-0.200	-0.380	0.000	-0.330	1.830	1.250
10	2.358	-0.340	-0.060	-0.160	-1.958	-0.080	0.280	-0.030	-0.170	-2.700	-0.580	1.030	2.250	-0.320	-0.200	0.050	1.030	0.920
11	1.974	-0.290	-1.000	-0.080	-0.684	-0.320	-0.080	0.240	0.160	-0.190	0.130	0.400	-0.340	-0.210	-0.340	-0.140	2.110	2.000
12	2.635	0.530	-0.280	-0.320	-2.885	0.190	0.330	0.250	-0.770	-2.410	0.630	0.850	0.930	-0.860	-0.280	-0.290	1.560	0.980
13	2.193	-0.320	0.270	-0.180	-2.143	0.480	1.040	0.120	-1.640	-1.310	0.070	0.490	0.750	-0.450	-0.470	-0.090	1.490	2.210
14	2.366	-0.470	1.950	0.540	-3.846	-0.370	-0.580	0.180	0.770	-0.030	-0.490	-0.440	0.960	0.220	0.340	-0.140	1.220	1.580
15	1.901	1.180	1.910	-0.820	-4.991	0.170	-0.050	-0.120	0.000	-0.420	-0.140	-0.120	0.680	0.060	-0.370	-0.200	1.350	1.540
16	0.919	-0.160	0.220	0.170	-0.979	1.340	0.990	-0.310	-2.020	-0.910	0.660	0.420	-0.170	-0.100	-0.090	-0.060	1.600	1.570
17	2.402	-2.270	0.860	1.140	-0.992	-0.010	0.030	0.300	-0.320	-1.570	-0.080	0.950	0.700	0.000	-0.010	-0.410	1.270	1.690
18	1.043	-0.420	2.040	0.750	-2.663	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.370	1.300
19	1.084	-0.030	-0.480	0.300	-0.574	1.090	0.060	0.130	-1.280	-0.010	-0.030	0.160	-0.120	-0.260	-0.070	-0.760	1.900	1.520
20	2.850	-0.250	-0.190	-0.150	-2.410	0.460	0.660	0.130	-1.250	-1.670	-1.370	1.260	1.780	0.020	-0.340	-0.150	1.520	1.050
21	1.232	-0.280	0.110	0.160	-1.062	1.010	1.060	0.290	-2.360	-1.150	-0.020	0.890	0.280	-0.270	-0.300	0.060	1.490	1.510

Abbreviations: MSR=mean squared residual; Phut=Pizza Hut; Dom=Dominos; Boys=Eagle Boys; Phav=Pizza Haven; ntop=number of toppings; bread=free bread; drink=free drinks; desrt=free desert; asc=alt-specific constant; $=price

Table A2a: Summary Statistics for Individual MNL Models - Flights

Effect	N	Mean	StdError	T-Stat	Skewness	StdError	Kurtosis	StdError
residual²	21	0.968	0.095	10.202	1.325	0.501	1.918	0.972
Time4hrs	21	0.334	0.095	3.525	0.862	0.501	0.076	0.972
Time5hrs	21	0.132	0.066	1.991	1.946	0.501	3.508	0.972
Time6hrs	21	-0.135	0.057	-2.362	-0.693	0.501	1.211	0.972
Time7hrs	21	-1.434	0.193	-7.429	-0.585	0.501	-0.811	0.972
Fare$350	21	1.403	0.134	10.439	-1.370	0.501	0.910	0.972
Fare$450	21	0.717	0.066	10.940	-0.799	0.501	1.403	0.972
Fare$550	21	-0.521	0.060	-8.751	0.648	0.501	-0.710	0.972
Fare$650	21	-1.598	0.137	-11.644	1.079	0.501	0.314	0.972
Qantas	21	0.076	0.105	0.719	2.687	0.501	9.752	0.972
Virgin Blue	21	0.200	0.086	2.312	1.946	0.501	3.783	0.972
JetStar	21	-0.258	0.161	-1.600	-2.646	0.501	6.726	0.972
AusAirlines	21	-0.017	0.107	-0.161	-1.710	0.501	8.365	0.972
FreqFlyer	21	0.010	0.020	0.480	0.655	0.501	0.498	0.972
# Stops	21	0.069	0.049	1.417	1.032	0.501	1.469	0.972
Drinks	21	-0.120	0.046	-2.585	-1.324	0.501	1.244	0.972
asc1	21	1.574	0.054	28.952	0.993	0.501	3.343	0.972
asc2	21	1.485	0.043	34.753	1.079	0.501	0.752	0.972

Table A2b: Summary Statistics for Individual MNL Models - Pizzas

Effect	N	Mean	StdError	T-Stat	Skewness	StdError	Kurtosis	StdError
Residual²	21	2.031	0.145	13.964	-0.225	0.501	-0.969	0.972
Pizza Hut	21	-0.158	0.160	-0.985	-0.554	0.501	2.891	0.972
Dominos	21	0.400	0.199	2.009	0.992	0.501	0.075	0.972
Eagle Boys	21	-0.062	0.186	-0.333	-2.092	0.501	7.472	0.972
Pizza Haven	21	-2.274	0.310	-7.346	-0.963	0.501	0.717	0.972
$12	21	0.309	0.117	2.638	0.386	0.501	-0.867	0.972
$14	21	0.347	0.098	3.557	0.019	0.501	-0.662	0.972
$16	21	0.113	0.056	2.026	0.168	0.501	0.274	0.972
$18	21	-0.769	0.195	-3.951	-0.224	0.501	-1.136	0.972
#toppings=1	21	-1.280	0.204	-6.264	-0.065	0.501	-0.859	0.972
#toppings=2	21	0.017	0.110	0.151	-0.604	0.501	2.058	0.972
#toppings=3	21	0.589	0.120	4.905	-0.196	0.501	-0.934	0.972
#toppings=4	21	0.675	0.154	4.382	0.449	0.501	-0.300	0.972
Free bread	21	-0.237	0.067	-3.515	-1.007	0.501	1.128	0.972
Free drinks	21	-0.234	0.058	-4.035	-0.066	0.501	0.426	0.972
Free desert	21	-0.181	0.049	-3.722	-0.830	0.501	0.840	0.972
Asc1	21	1.512	0.065	23.411	-0.119	0.501	0.090	0.972
Asc2	21	1.399	0.077	18.172	0.498	0.501	-0.140	0.972