Using fuzzy set theoretic techniques to identify preference rules from interactions in the linear model: an empirical study

Carl F. Mela[a,*], Donald R. Lehmann[b]

[a]*College of Business Administration, University of Notre Dame, 128 Hayes-Healy Center, Notre Dame, IN 46556-5641, USA*
[b]*Graduate School of Business, Columbia University, 213 Uris Hall, New York, NY 10027, USA*

Abstract

This paper seeks to establish a parametric linkage between fuzzy set theoretic techniques and commonly used preference formation rules in psychology and marketing. Such a linkage helps to benefit both fields. We accomplish this objective by using a linear model with interaction term which nests many common preference protocols; conjunction (fuzzy and), disjunction (fuzzy or), counterbalance (fuzzy xor) and linear compensatory. The resulting linear model with interactions can be employed when one has no a priori hypothesis about the individual's preference formation rule involved to determine the most likely preference rule or to test more formally the adequacy of a given rule. One illustrative application studies two-attribute decisions in six product categories and demonstrates differences in preference formation processes by product category. A second application demonstrates how fuzzy logical operators can be applied to situations involving more than two attributes.

Keywords: Empirical research; Management; Decision making

1. Introduction

There has been a call for research employing fuzzy set theoretic techniques in the social science [38]. We report one such application in the field of marketing. As Smithson [39] states, "many of the connections between fuzzy set concepts and mainstream research concerns in these areas (social sciences) have not been clearly drawn". The fundamental underpinning of many behavioral sciences is the linear model: we intend to draw that connection between fuzzy set theory and the mainstream linear model.

The hegemony of the linear model in marketing is one of the most enduring constants in applied social research. Linear models are especially prevalent in the analysis of attitude and choice [19, 24, 30, 35, 26]. Yet simple introspection suggests we rarely use the linear model as a process for coming to conclusions about the nature of choice. By employing fuzzy set theoretic operators, we intend to examine the link between preference structure and the parameters of the linear model with interactions. Applications of fuzzy set theory have found their way into finance [9], but are unknown to mainstream marketing. Finally, the axiomatic base of fuzzy operators used in set theory align well with much of the research to date in decision making.

* Corresponding author.

One of the reasons commonly cited for the use of the simple linear model is its predictive power [14, 16] and parsimony. This power may be due to positive correlation among attributes [13, 29]. Positive correlations are especially likely when attribute values are measured subjectively due to a halo effect [5] which makes even the weights of the linear model unimportant [4]. In addition, given the positive correlation among attributes, the inclusion of interactions becomes redundant. Consequently, many of these linear models omit first-order interactions terms. As an example, the use of the linear models *without* interaction terms has been especially prevalent in conjoint analysis [23].

Here, we extend the work of several researchers who have suggested the importance of interaction terms in modeling non-compensatory choice situations [26]. Furthermore, we show that given aggregation across heterogeneous (with respect to preference structure) consumers, the linear model will tend to increase in its predictive power due to the obscurement of individual level differences in preference formation strategies employed.

At one time, interest in the various types of non-compensatory choice models was high. Various models such as lexicographic, conjunctive and disjunctive were discussed in detail [6, 32, 44]. Farquhar and Rao [18] proposed an additional set of models, the balance and counter-balance. However, researchers suggest that such models as currently defined in marketing are rare in practice as they assume a very sharp non-fuzzy cutoff for acceptance of an attribute [32]. Recently, work on non-compensatory models has again emerged. For example. Gensch and Svestka [20] and Srinivasan [40] have proposed choice rules consisting of two phases, an elimination phase and a compensatory phase. Various choice and preference models such as EBA [42], and tree-based models such as EBT and HEM [43] have received considerable attention [31]. Similarly, nested-logit models [33] have been developed to explain non-linear choice. However, these models are relatively cumbersome to estimate and require a priori specification of the model form/choice rule [34]. The concept learning system [12] has been used as well to estimate logical choice rules. The model is difficult to estimate, however, and requires discrete attribute levels and is not based upon standard error theory. In this paper we suggest how estimating a linear model augmented with interaction terms can lead, through examination of its parameters, to the identification of preference rule used.

There is evidence that interaction terms have significance in non-compensatory choice. In one of the earliest works on the importance of interaction terms in non-compensatory choice, Bettman et al. [7], building on the work of Anderson [1], suggested a link exists between two-way interactions and the choice process. Anderson and Zalinski [3] further outlines two processes, adding and averaging, the combination of which produces a versatile model of preference formation. Anderson also suggests the model may have a cognitive interpretation. Further, he suggests fuzzy logic concepts may have an important application in understanding cognitive algebra. We shall presently formalize this conjecture. This paper extends this work and shows how the interaction term can be used to identify the choice process at the individual level and to segment consumers according to preference rule followed.

More specifically, we present a methodology that

- separates "and" (conjunctive) effects from "xor" (balance) and "or" (disjunctive) effects,
- measures degrees of "andness", "xorness", "orness", and linearity,
- uses the simple linear model augmented with interaction terms to estimate complex non-linear preference formation processes.

This paper is organized as follows: we first review the basic choice/preference models. Then, basic "and", "or", "xor", and negation functions will be introduced. Next a brief discussion of cognitive logic and its implications for the linear model is presented, followed by a discussion of estimation issues. We then present two illustrative applications of the approach. Finally, we suggest limitations and directions for future research.

2. Background

Several preference formation models, each representing multiple choice processes, have been proposed in the literature including both non-compensatory models and linear compensatory models. The degree to which a choice model is compensatory refers to the extent to which a deficiency in one attribute can be compensated for by another. The standard *linear* model is "partially" compensatory as a low rating in one attribute can be partially compensated for by higher ratings on other attributes. Several non-compensatory models have been suggested, the motivation for which are both the simplification of consumer choice strategy [32, 37].

A fully compensatory model as discussed in marketing and psychology (e.g. *disjunctive*) predicts that any attribute above a certain level will yield a fully favourable evaluation, regardless of the other attributes. The disjunctive model suggests that one selects an alternative by looking at its best attribute. For example, Einhorn [16] states, "In selecting players for a football team we might want someone who can kick *or* run *or* pass with a great deal of skill". In the simple linear model, this *cannot* be accounted for by a main effect alone. This is analogous to the fuzzy "or".

Non-compensatory models include the *conjunctive* and the *lexicographic* models [14, 16, 37]. In the conjunctive rule all attributes must exceed a certain level for the alternative to be evaluated favorably. For example, a computer may be more desirable only if it has both speed and memory, one or the other alone will not do. The lexicographic choice rule implies that an individual orders attributes by importance, selects the highest level on the best attribute and then moves down to the next attribute in the event of a tie. Although this process cannot be modeled exactly by a linear model [15] it should result in a strong main effect. The conjunctive model in psychology is analogous to the fuzzy "and".

A third class of choice processes, based upon the relative levels of the attributes also exists. For example, the balance model of Farquhar and Rao [18] suggests alternatives which have minimum dispersion across attributes are preferred. This model seems to apply to aesthetic judgements of beauty (i.e. a person's height and weight). The *exclusive or* (xor) strategy is the negation (opposite) of the balance (i.e. counterbalance) model of Farquhar and Rao [18], where maximum dispersion of attributes is most preferred. An example of this choice strategy is as follows: I may dislike loud, slow music, like soft, slow or loud, fast music, but dislike slow, loud music. Thus, as volume softens, and tempo increases, there is a crossover effect in preferences. This process is analogous to an exclusive or (xor) rule in fuzzy logic.

As we will show, these various rules can be captured via the addition of an interaction to the linear model and the use of product operators to interpret rules. *It is not clear a priori* whether the significance of the interactive effect in the linear model is due to (1) both attributes exceeding some level (conjunction), (2) an either/or effect, where if either attribute is present the product is rated highly (fully compensatory disjunctive vs. non-compensatory) or (3) any attribute exceeding some level (disjunctive model). We propose an empirical method to disentangle these effects.

3. Interpreting preference rules

3.1. A framework for the preference rules

In this section we show how a linear model with interactions can represent many of the choice rules discussed in the previous section. For ease of explanation on how the interactions capture preference formation, we first focus on examples using two binary attributes. In such situations, there are four distinct products possible. If we further consider preference to be binary, then there are 16 possible choice combinations of the four products (corresponding to choosing or not preferring the products). One can then enumerate (Table 1) all possible preference formation rules for the four products (conjunction is defined as "and" and disjunction is defined as "or"):

If these functions are defined in terms of the standard linear model with interactions, *it is possible to capture all 16 possible preference strategies.* The implication of this statement is that all possible preference formation rules can be represented paramorphically by the linear model. In addition, the continuous nature of the linear model with interactions enables us to cover convex combinations of basic preference strategies.

Often preference rules are not based on discrete attributes. In this section we examine the implica-

Table 1
List of dichotomous choice strategies

		Choice Strategy							
X_1	X_2	Independence	Conjunctive X_1, X_2	Conjunctive X_1, Not X_2	X_1 main	Conjunctive Not X_1, X_2	X_2 main	Xor	Disjunctive
0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1

		Choice Strategy							
X_1	X_2	Conjunctive Not X_1, not X_2	Balance not xor X_1, X_2	Not X_2 main	Disjunctive X_1, not X_2	Not X_1 main	Disjunctive not X_1, X_2	Disjunctive not X_1, not X_2	Independence
0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1

tions of using the concepts of fuzzy sets to describe non-discrete attribute preference rules. To capture the syntactic meaning of "and" and "or", the "and" function should be high when all attributes are high and the "or" function should be high when one of the attributes is high (we use the terminology "and" and conjunction as well as "or" and disjunction interchangeably).

3.2. *Fuzzy set theoretic techniques*

Rarely is a preference rule followed exactly. In this section we examine the implications of this using the concepts of fuzzy sets. To capture the syntactic meaning of "and" and "or", the "and" function should be high when all attributes are high and the "or" function should be high when one of the attributes is high (we use the terminology "and" and conjunction as well as "or" and disjunction interchangeably). To capture the syntactics of the "or" and "and" we use the following operationalizations of fuzzy conjunction and fuzzy disjunction (after [45, 47]):

$$\begin{aligned} &\textit{Conjunction} = x_1 x_2, \\ &\textit{Disjunction} = x_1 + x_2 - x_1 x_2, \end{aligned} \tag{1}$$

where $x_i x_j$ represent attribute levels. Attribute levels must be scaled or transformed to the interval between zero (minimum) and one (maximum or ideal) to maintain the idempotency property and interpretability of these operators (see Appendix A). The conjunction and disjunction surfaces based upon the linear model with interactions approximates the concave and convex surfaces suggested by Einhorn [16] for the conjunctive and disjunctive rules, respectively, suggesting a psychologically based precedent for the definitions in Eq. (1). Similar operators are defined for exclusive or and they are discussed in Appendix A (for example, we approximate the xor by $x_1 + x_2 - 2x_1 x_2$). Negation is defined as $1 - x$.

Mathematically, the logical functions should represent cognitive logic, and the functions should reduce to Boolean logic for dichotomous variables (see Appendix A). Along these lines, Zimmermann and Zysno [48] and Smithson [39] report encouraging experimental results regarding the explanatory power of the logical functions. Further, the functions have a strong axiomatic base. Moreover, the definitions of operators in fuzzy set theory are consistent with the traditional linear model with an interaction term.

3.3. Identifying the preference rules

To identify the various preference rules we need a flexible function such as Eq. (1) that allows us to use the product operators. This model is similar to the generalized adding model of Anderson and Zalinski [3] which nests many choice surfaces and is easily estimable. The model is also similar to Smithson [38] and is based upon the product operators determined earlier. The function is

$$\hat{y} = \beta_0 + \beta_1 \hat{x}_1 + \beta_2 \hat{x}_2 + \beta_3 \hat{x}_1 \hat{x}_2, \quad (2)$$

where $\hat{x}$ and $\hat{y}$ are rescaled into x and y by some conversion such as

$$\hat{x} = (x - x_{\min})/(x_{\max} - x_{\min}); \quad \hat{x} \in [0, 1].$$

For example, approximating the min–max operators defining conjunction disjunction and xor with product operators the conjunction of x_1 and x_2 can be modeled by $\beta_3 = 1$ and all other parameter equal to zero. This reproduces the preference rule for conjunction in Table 1. The conjunction of x_1 and not x_2 (or $1 - x_2$) is given by $x_1(1 - x_2)$, implying $\beta_1 = 1$, $\beta_3 = -1$, and all other parameters equal to zero. One can verify this set of parameters is also consistent with Table 1. Based upon the previous discussion of logic functions we know that parameter estimates map to the preference rules in Table 2 (with all points in between representing mixtures of these strategies).

The structure of the parameter table in Table 2 coincides with preference rules outlined in the preference table in Table 1. One interesting insight is that various preference strategies occur on a continuum captured by this model. Therefore, models such as a pure conjunction can be viewed as an extreme model, with actual processes representing various degrees of conjunction. Such an idea is not novel (cf. [16]), but many modelers have viewed the choice process as discrete, that is, individuals use pure and not mixed strategies. The model in Eq. (3) nests these various preference processes. It also reduces to the linear compensatory model when $\beta_2 = 0$, a variant of the Einhorn [16] conjunctive 1 when $\beta_1 = \beta_2 = 0$, $\beta_3 = 1$ and a variant of the Einhorn [16] disjunctive when $\beta_1 = \beta_2 = 1$, $\beta_3 = -1$. When parameters fall in between pure rules, it suggests a mixed strategy is being followed.

Table 2
Parameter table

Rule	Parameter			
	β_0	β_1	β_2	β_3
And(X_1, X_2)	0	0	0	1
And(X_1, not X_2)	0	1	0	−1
And(not X_1, X_2)	0	0	1	−1
And(not X_1, not X_2)	1	−1	−1	1
Or (X_1, X_2)	0	1	1	−1
Or(not X_1, X_2)	1	−1	0	1
Or(X_1, not X_2)	1	0	−1	1
Or(not X_1, not X_2)	1	0	0	−1
Xor (X_1, X_2), Xor (not X_1, not X_2)	0	1	1	−2
Xor(not X_1, X_2), (Xor (X_1, not X_2)	1	−1	−1	2
Main effect (X_1)	0	1	0	0
Main effect (not X_1)	1	−1	0	0
Main effect (X_2)	0	0	1	0
Main effect (not X_2)	1	0	−1	0
Independence (ideal)	1	0	0	0
Independence (substandard)	0	0	0	0

The critical difference between a fuzzy conjunction and disjunction and a standard conjunction and disjunction is that the "cutoff" levels for the conjunction and disjunction are fuzzy rather than discrete. More formally, consider Fig. 1, a graph of conjunction for two attributes. Standard conjunction (Fig. 1(a)) specifies cutoff levels for attributes, x_1^* and x_2^*. The acceptance region for the product is ($x_1 > x_1^*$ (e.g. 0.5), $x_2 > x_2^*$ (e.g. 0.5)). Any alternative that meets this criteria is completely acceptable. A fuzzy conjunction (Fig. 1(b) and (b′)) relaxes this assumption; rather than a fixed cutoff, there exists a gradual cutoff for which the alternative is somewhere between completely acceptable and completely unacceptable. The fuzzy conjunction reduces to standard conjunction when the range of the cutoff reduces to zero.

Mixed choice strategies can be captured via a convex mixture set of a logical "and", a logical "or", and a logical "xor", including the linear model without interactions. The definitions suggest that "and" and "or" represent *equal and opposite* departures from linearity as the interpretation of an equal amount of cognitive "and" and cognitive "or"

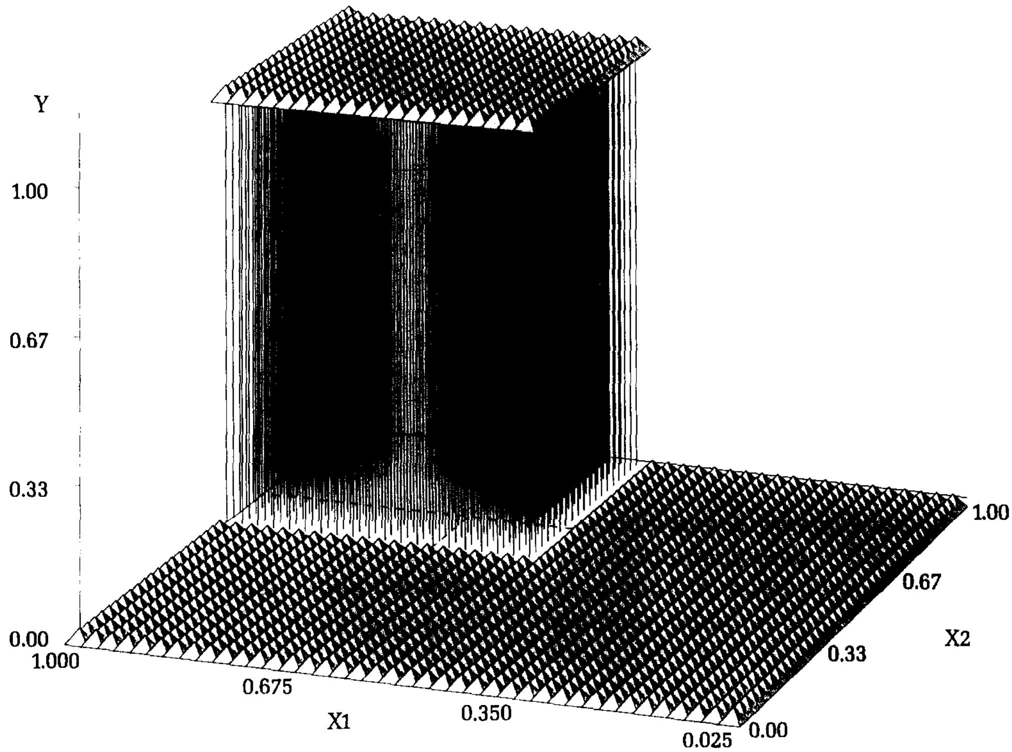

Fig. 1(a). Fixed conjunction.

(neither "and" or "or" dominate) is simple linearity [38]. The linear model is partly compensatory and partly non-compensatory to the extent it lies between disjunctive (completely compensatory) and conjunctive (completely non-compensatory). It is possible to disentangle "or" and "and" effects via the coefficient on the interaction term. A positive coefficient is more "and" than "or" and a negative coefficient is more "or" than "and".

To identify the rule used, one can compare the parameter vector listed in Table 2 to the parameter vector estimated from running the regression in Eq. (1). Further, one can run statistical tests to test hypothesis about whether or not a specific rule is being followed. There exists a statistic to test a set of specific hypothesis about parameter vectors, $\boldsymbol{\beta} = \boldsymbol{r}$. The statistic is a standard distance metric, distributed $F_{q,n-k}$, which allows for testing of specific rules.[1] The numerator of the F statistic can be written $Q = (\boldsymbol{b} - \boldsymbol{r})'((X'X)')^{-1}(\boldsymbol{b} - \boldsymbol{r})/q'$ and the denominator is the mean squared error, $\boldsymbol{e}'\boldsymbol{e}/n$ [25]. In this instance, $\boldsymbol{r}$ is simply the vector representing a given choice rule. If one has a prior theory concerning the rule followed, this statistic provides a test of the theory. We can also use this statistic as our distance metric and classify subjects according to the decision rule they most closely follow, including priors if they are available.

[1] Assume one wishes to test a set of specific hypothesis about parameter vectors, $\boldsymbol{R\beta} = \boldsymbol{r}$. The numerator of the F statistic can be written $Q = (\boldsymbol{Rb} - \boldsymbol{r})'(\boldsymbol{R}(X'X)\boldsymbol{R}')^{-1}(\boldsymbol{Rb} - \boldsymbol{r})/q$ (where $\boldsymbol{R}$ is dimensioned $q \times k$, in our case, a 3×3 identity matrix for the three parameters) and the denominator is the mean squared error, $\boldsymbol{e}'\boldsymbol{e}/n$) [25]. In this instance, $\boldsymbol{R} = \boldsymbol{1}$ the matrix of diagonals as we are only testing for one particular rule, and $\boldsymbol{r}$ is simply the vector representing a given choice rule. Thus, the $\boldsymbol{R}$'s drop out of this expression, yielding the test $Q = (\boldsymbol{b} - \boldsymbol{r})'(X'X)^{-1}(\boldsymbol{b} - \boldsymbol{r})$.

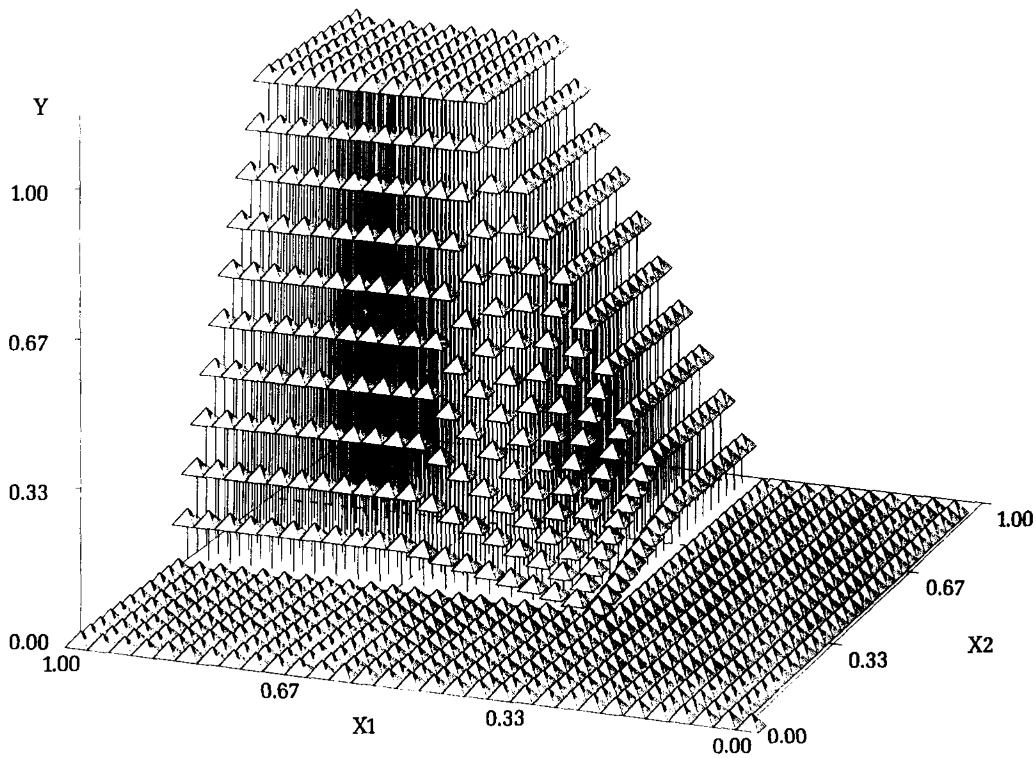

Fig. 1(b). Fuzzy conjunction.

3.4. Attribute valuations

We have assumed preference is linear in attributes over the range $x_{\min}$, $x_{\max}$. We may relax the assumption at the cost of several parameters. The preference model now becomes, similar to Elrod [17] with an interaction added:

$$y = \beta_1 f(x_1) + \beta_2 f(x_2) + \beta_3 f(x_1) f(x_2), \tag{3}$$

where $x \in [x_{\min}, x_{\max}]$ is replaced by a mapping of the original attribute space to a revised attribute preference space. Note that a constant term is not included in this model, as there are four cells and four parameters and the model would be fully saturated. The intercept allowed for negation in the linear model; here negation is incorporated directly into $f(x)$. For example, not $\text{and}(x_1, x_2) = 1 - \text{and}(x_1, x_2) = \text{or}(\text{not}\, x_1, \text{not}\ x_2) = \text{or}(f(x_1), f(x_2))$.

Anderson et al. [2] suggest standardizing attributes in order to facilitate parameter interpretation by scaling the attributes between zero and one. Anderson and Zalinski [3] note that such a conversion may confound attribute ranges and weights, however. The proposed alternative is said to be difficult to estimate in practice. We therefore employ the Anderson [2] standardization given in Eq. (2). Specifically, if one believes the attribute function is linear on the closed interval [0, 1] one can simply convert the raw attributes to a linear scale $((x - x_{\min})/(x_{\max} - x_{\min}))$. Thus, the acceptance region for the conjunction or disjunction is allowed to become fuzzy rather than crisp relaxing the severe assumption to crisp cutoff values for attribute acceptance. Other transformations, $f(x)$, include the logistic (for estimating the cutoff regions) and mean squared deviations for ideal point models (attributes such as saltiness). For example, the logistic

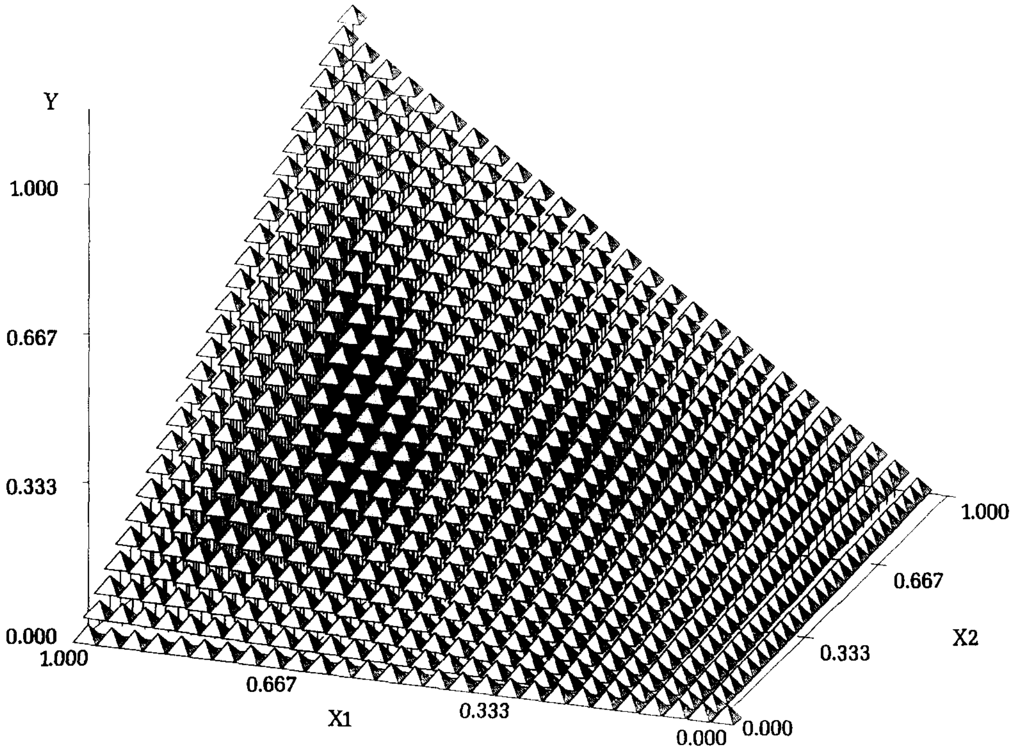

Fig. 1(b'). Fuzzy conjunction.

transformation is given by

$$f(x_i) = \frac{\exp(\beta x_{ri} - k)}{1 + \exp(\beta x_{ri} - k)}$$

where k approximates the beginning of the consumer's fuzzy cutoff point (below which the alternative is unacceptable), and β approximates (1 − end of fuzzy cutoff)/(end of fuzzy cutoff − beginning of fuzzy cutoff). As β approaches infinity, the model reduces to a standard non-fuzzy conjunctive or disjunctive model. Whether the linear or logistic model provides a better fit is an empirical issue. However, the cost of additional parameters and non-linear estimation (e.g. non-linear least squares or maximum likelihood) makes adding the attribute transformation step somewhat unwieldy.

In sum, the parameters of (1) can be estimated via at least three transformations of attributes to the [0, 1] scale – linear. logistic, and ideal point. Alternatively, 0–1 attribute level ratings can be obtained directly from subjects (attribute perceptions). The choice of the best approach undoubtedly depends on both individual and situational characteristics and is not the focus of this paper.

3.5. Segment level estimation

Interpreting individual results can be cumbersome. In addition, aggregate level estimation can result in aggregation bias. To simplify interpretation and increase robustness of the estimates (at the possible cost of aggregation bias), we can cluster respondents into segments by their choice processes. This can be accomplished by estimating individual models and clustering individuals by standard clustering routines. Alternatively, if the number of observations per individual is too limited for stable estimates, at the cost of more complex estimation one can perform latent class analysis by estimating mixtures of regressions.

One can also relate memberships to individual characteristics (e.g. demographics) in order to assist in targeting particular choice rule segments. To do this, one must regress membership values on a set of characteristics [10].

3.6. Extensions to more than two attributes

To extend the methodology to more than two variables is conceptually straightforward although the "accounting" is difficult. Two-way interactions have already been considered. Three- and higher-way interactions can be products of interactions between two way interactions, such as $((x_1$ and $x_2)$ and $x_3) = (x_2 x_2)$ and $x_3 = x_1 x_2 x_3$. One way to proceed is by first hypothesizing the various choice rules, and then applying the operators. For example, assume a choice of soda is a function of [(promotion or display) and (diet)]. The parametric form of this model is [(promotion + display − promotion * display) * diet] = promotion * diet + display * diet − promotion * display * diet. A serious limitation in extending this approach to multiple attributes is the number of profiles that must be given to an individual for a given product in order to test for all or some of the potential interactions at the individual level.

Therefore, estimation of higher-order models follows a different procedure. Due to the higher number of parameters associated with the higher-order interaction terms, parameter estimates based upon individual level estimation becomes increasingly unstable. For example, two attributes require estimating four parameters. With three attributes, one must estimate eight parameters. Four attributes result in fifteen parameters and five attributes result in twenty-six parameters, and so on. With twenty or so profiles (observations/subject), degrees of freedom and consequently the reliability of estimates at the individual level are adversely impacted. On the other hand, estimation at the aggregate level suffers from aggregation bias, which, as we suggest, may obscure conjunction and disjunction to some degree.

A similar problem is faced in the logit choice modeling literature and was resolved by Kamakura and Russell [27]. Essentially, they employ a latent clustering methodology that assigns individuals to segments based on parameter vectors. They estimate a mixture of logits, in our case we estimate a mixture of normals. The result of the procedure, then, is several segments, each with its own parameter vector (or equivalently, choice rule), with individuals having certain probabilities of membership in each segment.

We start by assuming a segment, j, has the fuzzy choice rule of n attributes given by the higher-order extension of (2),

$$y_j = \beta_{j0} + \sum_{k=1,\ldots,n} \beta_{jk} x_k + \sum_{\substack{l,m=1,\ldots,n \\ l \neq m}} \beta_{jkl} x_k x_l + \cdots + \beta_{jlm\ldots n} x_k x_l \ldots x_n + \varepsilon_j \quad (4)$$

and $\varepsilon_j \sim \text{Normal}(0, \sigma_j)$. The likelihood function for this segment across time periods t can be written as

$$L_j = \prod_j \prod_t \phi\left(\frac{y - \boldsymbol{\beta}_j \boldsymbol{x}}{\sigma_j}\right). \quad (5)$$

The aggregate level likelihood is a weighted sum of the segment level likelihoods with the weights, w, representing the segment sizes:

$$L = \sum_{j=1}^{J} w_j L_j; \quad \sum_{j=1}^{J} w_j = 1; \quad w_j \in [0, 1]. \quad (6)$$

The number of segments is determined by the Beysian information criterion (BIC) for testing competing models, where $BIC = LL - (K/2)\log(n)$ where LL is the log likelihood value, k is the number of parameters and n is the number of observations. Essentially the BIC trade-off fit and increasing complexity which results from increased parameters. Finally, two minor points need to be raised regarding estimation. The segment standard deviations need to be constrained to be positive, and the segment weights must be positive and sum to one. Both constraints may be met during estimation by first setting the segment standard deviations to be a squared function of a parameter to be estimated, $\sigma_j = \sqrt{\hat{\sigma}_j^2}$ and second setting $w_j = \exp(\gamma_j)/\sum\exp(\gamma_j)$ and estimating the γ. Finally, the posterior probability of an individual belonging to a given segment is given by

$$p_{ij} = \frac{w_i L_{ij}}{\sum_{j=1}^{J} w_j L_{ij}}. \quad (7)$$

In sum, higher-order interactions and more complex choice rules can be estimated at the segment level via maximum likelihood.

4. Illustrative examples

We present two illustrative examples of the methodology. The first example uses two attributes and is presented in order to (i) ascertain the face validity of the model and (ii) demonstrate application of this technology to inferring choice protocol, studying intra-subject consistency in choice.

The second example is presented to illustrate estimation of the model with more attributes (three) to show how latent class models can be used to uncover segment of model users when data does not permit individual level estimation.

4.1. Study 1

A convenience sample of 30 MBA students evaluated twenty alternatives in each of six categories. The categories (and attributes) were stereos (speaker size, receiver size), television shows (laughter, horror), apartments (size, commute), computers (memory, speed), gymnasiums (number of weight machines, number of aerobic machines), and college applicants (GMAT, GPA). These categories were chosen because of their relevance to the subject population (based upon pretest results) and because the rule followed might be expected to differ by category. Given the task employed, we expected the linear model (the model with low weights on the interaction term) to perform well. However, we do expect variation across the categories. Specifically, we expect stereos to produce more balance rules (matching speaker and receiver sizes), and television shows to produce more disjunctive rules (since humor and horror may interfere with each other). The overall task took 45 minutes.

The twenty product profiles in each category were developed by choosing ranges of the attributes that approximately corresponded to the minimum and maximum levels beyond which stated preference levels for a given attribute did not vary substantially in a pre-test. The attributes were presented in a random order to the subjects such that the inter-attribute correlation was zero.

Subjects first provided evaluation of the twenty alternatives for each of the six categories on a 0–100 scale. We then asked subjects for retrospective verbal protocols of their choice process.

4.2. Individual level results

Following Farquhar and Rao [18], we first report the results of typical individuals and then present general results of the model. For example, subject 2's self-explicated preference for stereos is as follows (parenthesis added): "It's (the stereo system) either portable (small) or solid (large) – any exaggerated mix (non-match between speaker size and stereo size) doesn't work". This seems to suggest a balance model. The parameter estimates for this person are (1.11, − 0.81, − 1.30, 2.18), quite similar to the pure balance model which is given by (1, − 1, − 1, 2). Thus, the model seems to recover the choice rule the subject believed he was using.

Subject 11's self-explicated preference for laughter and horror in television shows is as follows, "I like both horror and comedy movies, but would not seriously like to see both in the same movie". Thus, we have an xor. The parameter vector is (− 0.26, 0.69, 0.93, − 1.56) which is close to the pure xor, given by (0, 1, 1, − 2). This subject is apparently pursuing a strategy that is more xor than "or".

Subject 19's self-explicated preference for computers is as follows, "If either attribute was below a certain level, the product was unacceptable, if both were above the level I wanted the most of each". Thus, we have an "and". The parameter estimates are (0.07, 0.04, 0.02, 1.00), close to the pure conjunction which is given by (0, 0, 0, 1).

Subject 9's self-explicated preference for applicants to MBA programs is "did not like low (test scores)–low(GPA)". This subject thinks they are pursuing an "or" strategy. The subject's parameter

Table 3
Category choice rules

Category	Parameter				Rule
	β_0	β_1	β_2	β_3	
Computers	0.08	0.24	0.15	0.63	Mostly conjunctive
Stereos	0.41	0.17	− 0.29	0.46	Mixed conjunction/independence
Apartments	0.28	− 0.25	0.77	− 0.38	Linear, time negatively scaled
Gymnasiums	0.13	0.25	0.21	0.53	Mixed conjunction/linear
TV shows	0.09	0.15	0.89	− 0.57	Conjunction of X_1, not X_2
MBA applicants	− 0.12	0.53	0.46	0.18	Linear

vector is (− 0.01, 1.12, 0.96, − 0.97), compared with the pure conjunction given by (0, 1, 1, − 1).

Last, as an example of a linear compensatory strategy, subject 16 said "I gave more importance to time to school rather than size, but looked for a reasonable size as well". Assuming that distance is twice as important as size, and that greater distance is negatively scaled, the subject's self explicated rule suggests the parameter estimates would be (0.67, − 0.67, 0.33, 0.00). The estimated parameter vector is given by (0.73, − 0.68, 0.37, − 0.05). Thus, the method seems to recover choice rules which match subject's perceptions of their choice process. We next ascertained how well the inferred choice rule correlated with self-explicated choice rules. Since many rules fall on a logical continuum, we coded the self-explicated rules on a 1–4 scale with 1 representing conjunction, 2 representing linear, 3 representing disjunction and 4 representing xor. We used as the inferred choice rule measure, β_3 in (1), which runs from (− 2, 1] along the rule continuum from xor to "and", respectively, and is therefore a measure of the disjunctiveness of the rule used.

The inferred scale should correlate negatively with coding of the self-explicated rules (as self-explicated conjunction is scaled 1 – high conjunction/low disjunction 4 – low conjunction/high disjunction. The two scales correlate negatively (− 0.26, $p = 0.0006$) providing both face validity to the results and some more formal evidence that subjects were aware of the rule they were following.

One way of looking at the markets in aggregate is through the average parameters of the model. For each category we average parameters across individuals, producing Table 3. Interestingly, the aggregate results suggest that hybrid, and not pure, choice rules are being followed. With the exception of apparently linear choice for computers and MBA applicants, the choice patterns tend to be mixed. However, given the possibility of aggregation obscuring heterogeneity in choice patterns, we attempted to segment subjects by the choice rules followed for each product category.

4.3. Segmentation results

Individuals were clustered on the basis of Euclidean distance between their parameter vectors. The number of segments was determined by the pseudo-F statistic which is an overall measure of between/within variance that is adjusted for the number of parameters [11]. The statistic is given by $(R^2/(c-1))/(1-R^2/(n-c))$ where c is the number of clusters and n is the number of observations. A second criteria for selecting the number of clusters was also applied, which consisted of reasonable segment sizes and interpretability. If additional insight into the market was obtained by adding or subtracting a segment and the pseudo-F statistic was within 1 of the optimal and an added segment had only one or two individuals, we chose the more parsimonious or interpretable solution. This occurred once, for stereos, where the F statistic changed only marginally (1.0) and the added cluster only had two persons. Table 4 presents the overall

Table 4
Cluster analysis of rules followed by category

Category	R^2	Pseudo-F	Segment-Persons	β_0 (S.D.)	β_1 (S.D.)	β_2 (S.D.)	β_3 (S.D.)	Rule
Computers								
Linear in attributes	0.38	35.87	1 13	0.14 (0.18)	– 0.03 (0.17)	– 0.06 (0.17)	1.09 (0.19)	Conjunction
			2 17	0.04 (0.21)	0.44 (0.28)	0.31 (0.31)	0.27 (0.25)	Mostly linear
Attribute perception	0.41	26.95	1 14		0.08 (0.12)	0.09 (0.28)	0.89 (0.28)	Conjunction
			2 15		0.61 (0.30)	0.25 (0.30)	0.17 (0.29)	Mostly linear[a]
Stereos								
Linear in attributes	0.37	22.92	1 10	0.73 (0.33)	– 0.35 (0.32)	– 0.73 (0.48)	1.29 (0.57)	Mostly balance[b]
			2 19	0.24 (0.30)	0.31 (0.38)	– 0.04 (0.39)	0.20 (0.49)	Mixed main, conjunction[c]
Attribute perception	0.53	32.14	1 6		0.92 (0.46)	0.95 (0.59)	– 1.56 (0.63)	Mixed disjunction, Xor
			2 23		0.40 (0.30)	0.36 (0.33)	0.14 (0.47)	Mostly linear
Apartments								
Linear in attributes	0.34	41.60	1 14	0.61 (0.22)	– 0.53 (0.26)	0.42 (0.24)	– 0.01 (0.34)	Linear[d]
			2 16	0.01 (0.19)	– 0.01 (0.19)	1.09 (0.20)	– 0.71 (0.42)	Conjunction[e]
Attribute perception	0.43	27.30	1 14		0.39 (0.39)	0.56 (0.23)	– 0.02 (0.28)	Linear
			2 16		0.10 (0.16)	0.18 (0.25)	0.73 (0.28)	Mostly conjunction
Health clubs								
Linear in attributes	0.38	30.99	1 15	0.04 (0.20)	0.42 (0.31)	0.45 (0.32)	0.13 (0.22)	Linear
			2 15	0.21 (0.09)	0.09 (0.26)	– 0.04 (0.16)	0.93 (0.28)	Conjunction
Attribute perception	0.69	36.38	1 13		0.58 (0.26)	0.43 (0.34)	– 0.08 (0.27)	Linear
			2 15		0.23 (0.11)	0.19 (0.17)	0.64 (0.27)	Mostly conjunction
			3 1		1.09	1.16	– 2.39	Xor
Television shows								
Linear in attributes	0.56	29.98	1 6	0.48 (0.21)	– 0.36 (0.29)	0.45 (0.24)	– 0.02 (0.23)	Linear[f]
			2 5	– 0.20 (0.05)	0.95 (0.26)	1.18 (0.11)	– 1.23 (0.21)	Mostly disjunction
			3 19	0.06 (0.13)	0.10 (0.22)	0.94 (0.16)	– 0.57 (0.32)	Mostly conjunction[g]
Attribute perception	0.64	29.24	1 11		0.22 (0.25)	0.64 (0.28)	0.12 (0.16)	Linear
			2 9		0.07 (0.19)	0.34 (0.22)	0.59 (0.28)	Mostly conjunction
			3 9		0.56 (0.21)	0.86 (0.86)	– 0.61 (0.24)	Mostly disjunction
MBA admissions								
Linear in attributes	0.58	45.42	1 10	0.02 (0.16)	0.18 (0.18)	0.18 (0.12)	0.81 (0.24)	Conjunction
			2 6	– 0.33 (0.17)	1.01 (0.17)	0.90 (0.28)	– 0.66 (0.31)	Disjunction
			3 14	0.00 (0.17)	0.47 (0.16)	0.51 (0.24)	0.06 (0.20)	Linear
Attribute perception	0.54	69.57	1 17		0.51 (0.19)	0.49 (0.18)	– 0.05 (0.25)	Linear
			2 13		0.19 (0.18)	0.07 (0.13)	0.87 (0.24)	Conjunction

[a] Speed is weighted greater than memory.
[b] Speaker size weighted more than stereo size.
[c] Smaller stereo size is preferred.
[d] Commute time is negatively scaled.
[e] Large size, not distant.
[f] Horror is weighted negatively, i.e. not(horror).
[g] Conjunction of laughter and not horror.

results for the linear in attributes model (Eq. (2)) and the attribute perception model (Section 3.4).

The most striking result one notices from Table 4 is how the markets segment. Interestingly, mixed strategy segments rarely appear. The segments generally represent pure, rather than mixed, strategies suggesting segmentation by choice rule may be a very effective way to divide the market. Also, aggregation tends to drive the overall result to linear, and that most categories have a conjunctive and a linear segment. In the MBA admissions category, we notice conjunctive and disjunctive segments canceling each other out such that the interaction parameter tends toward zero at the aggregate level.

A question of interest centers about the consistency of choice strategy across categories. Earlier findings [28] suggest that innovators differ across categories. The parallel here is that conjunctive or other rule users may differ across category. To examine this issue, we noted the aggregate percentage, p, of those using conjunctive choice strategies. Assuming that consumers randomly use the conjunctive strategy with probability, p, a certain number of persons will use a conjunctive rule across all categories and a certain percentage use conjunctions in all but one category and so on. Since five out of the six categories contained a conjunctive segment, we calculated the aggregate probability of a conjunction across the five categories, which was 48%. Based upon this, we would expect to observe the number of persons using (5 conjunctions, 4 conjunctions, ..., 0 conjunctions) to be (0.8, 4.1, 9.0, 9.7, 5.2, 1.1). If individuals do differ in their propensity to use conjunction, we would expect a more extreme distribution, with a large number of individuals using conjunction in all five categories and a large number using not using conjunction in any category. We observed (2, 2, 12, 7, 4, 3) for the frequencies which produces a chi-square of 4 degrees of freedom, $\chi = 8.53$ ($p > 0.10$). As this does not differ significantly from chance results, .we conclude that choice strategy tends to be *category and individual specific rather than individual based.*

4.4. *Study 2*

Study two was undertaken to illustrate the application of the preference formation model to higher numbers of attributes. The data from study 2 was taken from a conjoint task of car tires reported in Green [22]. In this study 250 respondents were required to evaluate 25 profiles in a Graeco-Lattin design involving five levels of tread mileage (30 000, 40 000, 50 000, 60 000, and 70 000), five price level ($40, $55, $70, $85 and $100), and five distances to tire outlet (10 min, 20 min, 30 min, 40 min, and 50 min). In addition, there were five brands, *A–E*. The profiles were rated on a scale of 0–5. We concentrated on three attributes, mileage, price and distance, as these were easily convertible to a 0–1 scale and because the brand was insignificant at the aggregate level in a main effects conjoint analysis. We focus on estimating latent segments in this study.

Table 5
BIC table

Segments	Log likelihood	Parameters	*BIC*
1	− 4739.6	9	− 4778.8
2	− 4602.9	19	− 4685.7
3	− 4526.7	29	− 4653.1
4	− 4496.7	39	− 4666.4
5	− 4496.4	49	− 4710.0

4.5. *Results*

The parameters once again all are in the expected ranges. The optimal number of segments is three (see Table 5). We again employed the latent class segmentation approach using (4) and (5). Based on the *BIC* criteria, the three-segment solution is best. The results for the three-segment solution are presented in Table 6.

The three segments can be characterized by looking at the significant terms in each segment. Segment 1, 41% of the market is a mixed segment, price is a linear attribute but there is also a conjunction of miles and price. A mixed strategy like this with price negatively scaled would be given by $\frac{1}{2}(1 - \text{price}) + \frac{1}{2}(\text{miles} * (1 - \text{price})) = \frac{1}{2} - \frac{1}{2}\text{price} + \frac{1}{2}\text{miles} - \frac{1}{2}\text{miles} * \text{price}$, or in vector form, (0.5, − 0.5, 0.5, − 0.5). The actual parameters are (0.5, 0.4, − 0.5, − 0.3).

Segment 2 is similar to segment 1 in choice rule but not attributes. It is a mixed segment, price is a linear attribute but there is also a conjunction of miles and time to store. The parameter vector of such a strategy is $\frac{1}{2}(1 - \text{price}) + \frac{1}{2}(\text{miles} * (1 - \text{time}))$, or (0.5, − 0.5, 0.5, − 0.5). The actual vector is (0.3, − 0.6, 0.4, − 0.8). Interestingly, the three-way interaction term is nearly significant, offering evidence of a slight tendency toward a conjunction of high miles, low price, and low time to store. Were this the strategy, the choice rule would be $0.5(1 - \text{price}) + 0.5(1 - \text{price}) * (1 - \text{time}) * \text{miles} = 0.5 - 0.5 * \text{price} + 0.5 * \text{miles} - 0.5 * \text{miles} * \text{price} - 0.5 * \text{miles} * \text{time} + 0.5 * \text{miles} * \text{price} * \text{time}$, or (0.5, − 0.5, 0.5, − 0.5, − 0.5, 0.5). The actual vector is (0.3, − 0.6, 0.4, − 0.1, − 0.6, 0.5). The miles $*$ price term is not significant (parameter = 0.13, $t = -0.9$), preventing us from

Table 6
Three segment solution

Parameter	Value	*T* statistic
Segment 1		
Intercept	0.53	15.0
Mile	0.37	5.6
Time	− 0.03	− 0.5
Price	− 0.47	− 5.3
*Price * Mile*	− 0.32	− 3.2
Price * Time	0.02	0.2
Mile * Time	0.08	0.4
Price * Mile * Time	− 0.23	− 1.1
Standard deviation[2]	0.59	66.2
Size (γ)	− 0.07	− 0.3
Segment 2		
Intercept	0.32	6.2
Mile	0.41	4.2
Time	0.13	1.3
Price	− 0.61	− 4.5
Price * Mile	− 0.13	− 0.9
Price * Time	0.20	1.4
*Mile * Time*	− 0.82	− 3.1
Price * Mile * Time	0.52	1.8
Standard deviation[2]	0.56	96.6
Size (γ)	− 0.11	− 4.2
Segment 3		
Intercept	0.53	11.9
Mile	0.22	2.4
Time	− 0.08	− 0.8
Price	− 0.07	− 0.6
Price * Mile	− 0.25	− 1.8
Price * Time	− 0.14	− 1.1
Mile * Time	− 0.01	− 0.3
Price * Mile * Time	0.22	0.7
Standard deviation[2]	0.69	

Note: Significant attributes are listed in *italics* ($p < 0.05$ one tail test).

categorizing this as a highly conjunctive segment, although the term is correctly signed. This segment is small, representing about 14% of the market. Segment 3 respondents are fairly indifferent between brands with a tendency towards preferring brands with high mileage. The third segment is 45% of the market. The evidence, then, for use of conjunction as a simplifying strategy is somewhat limited in this category as the evidence suggests the use of two-way configural evaluations but not the complete three-way conjunctions. The three-way conjunction was only nearly significant and only for 14% of the market.

We can also briefly touch upon the one–four-segment solutions to see which segments appear first. The one-segment solution is largely similar to the segment 1 described in the three-segment solution. The two-segment solution is similar to segments 1 and 3 in the three-segment solution, and this is not surprising as these are the two largest segments. The three-segment solution adds the more conjunctive segment, and the four-segment solution splits the largest segment in the three-segment solution, the main effect for miles segment, into two. The new segments become main effect for miles, and an "indifferent" segment (no terms except intercept significant). The above analysis can be very helpful for targeting specific market segments with products or media campaigns based upon how individuals make decisions.

5. Summary

Fuzzy set theory provides insight into the linear preference formation model by clarifying the meaning behind the interaction term in the standard linear model. We also hope that the axiomatic base of the linear model provides insight into fuzzy set theoretic operators. We find strong empirical support for the use of fuzzy operators in the field of human behavior. Essentially, we estimated the standard linear model with interactions (after transforming the attributes to a 0–1 scale) in order to infer the consumer choice rule used. We show that this model can describe both compensatory and non-compensatory choice rules. As this form of choice rule inferencing uncovers choice strategies, it can be a useful form of decompositional analysis.

We estimated the model on 30 individuals in six product categories. The model's inferred choice protocol seemed to reflect the subject's stated choice rules fairly well. However, we also demonstrate that heterogeneity in terms of some individuals using conjunction with other individuals using disjunction or xor *will drive the results toward linear compensatory in the aggregate.* We therefore clustered individuals by their choice rules, and found that segments of fairly distinct choice strategies tend to emerge. In addition, we analyzed individuals to see if they tended to use the same choice rules across categories, and found no evidence of this. We also outlined an approach to estimate the model for higher numbers of attributes. In the second study again we find evidence that aggregation can cover up conjunctive processing strategies.

Of course, the model, although very flexible, has a number of limitations. First, the attributes must be ratio scaled and scaled between one and zero for the product operators to satisfy the basic logical axioms and be ratio scaled. In addition, in OLS it is possible to obtain estimates that yield values for the dependent variable slightly out of the [0, 1] interval. Further, the normality assumption is violated if one applies the constraint that main effect parameters should lie in [1, − 1] and the interaction term is also bounded in the interval [1, − 2]. Also, the linear attribute and logistic attribute models are not nested, making comparison somewhat difficult.

Here, choice rule inferencing was done on preference data, and like conjoint analysis, predicts value, not choice. In this sense, the analysis is very similar to most conjoint analysis. In that sense it adds an extra dimension to decompositional analysis of preference. The model may also be applied to choice, with the dependent variable being dichotomous in [0, 1] rather than continuous. The model can also be extended to the logit choice model; for example, an interaction going to positive infinity with no main effects would be indicative of a pure conjunction.

Third, the interpretation of conjunction is largely a mathematical definition, although we argue there is a cognitive interpretation as well. One might also view the process of conjunction as complements and disjunction as substitutes, or perhaps offer additional interpretations. In our view, all interpretations are consistent as they are paramorphically identical.

Substantive extensions regarding choice rule inferencing may be related to factors that affect the choice strategy (e.g. comparability of alternatives, involvement, decision complexity, familiarity, group process, etc.) which can add valuable marketing insight and offer a promising next step. By using this method, research could examine which factors affect the compensatoriness of choice. Alternatively, one could also use this model to study choice dynamics. One can also see how the choice rules change when new attributes or alternatives are introduced. The method can also be incorporated into choice models for determining optimal product design as it uncovers specific sets of combinations of attributes consumers prefer. Methodologically, we would like to see the model extended to discrete choice applications.

In this paper we have attempted to unify various choice and preference rules into a single, straightforward framework, the linear model with interactions, thereby enhancing the insights gained from the linear model. Furthermore, all the choice rules are nested within this single framework, making the tool a flexible model for research into consumer preferences. We also believe the model has clarified the interpretation of the interactions in linear models. Finally, we have attempted to tie together the mathematical and cognitive interpretations by discussing how parameters and model forms relate directly to the nature of preference formation.

In closing, we believe that fuzzy set theory shows some promise as a useful tool in market research. Choice rule inferencing highlights the explanatory power of the linear model. We believe that the robustness of the linear model not only lies in its predictive ability, but in its explanatory power as well.

Appendix A. Logic definitions

We define multiplicative "and" here as and = x_1x_2. The corresponding "or" = $x_1 + x_2 - x_1x_2$. The limitations of the product rule are that it must be a ratio scale bound between zero and one [21]. An intuitive explanation suggests that the and (intersection) of two membership values cannot exceed the value (size) of either of the memberships. Since multiplication and division can only be performed on ratio scales, we are thus limited in the use of ratio numbers for multiplicative fuzzy set logic functions.

The xor function is defined by DeMorgan's theorem as xor = (not and) and or. We simplify this as follows, (not and) and or = (1 − and) and or = or − (and and or) which is approximately (or − and) or $x_1 + x_2 - 2x_1x_2$. To simplify, we need to approximate that and and or = and. We assume this since "and" is more restrictive than "or", and we approximate (and) and (or) by min(and, or) instead of and times or. Recall, "and" is an intersection, and since and is a subset of or, the intersection of a set is its subset. This is not quite true for the product operator (due to lack of idempotency) unless we modify the operator definition to state that x and $x = x$ instead of x and $x = x^2$. Although this is an approximation of DeMorgan's theorem, it provides for more intuitive results. Therefore, an xor = − and $= x_1 + x_2 - x_1x_2 - x_1x_2 = x_1 + x_2 - 2x_1x_2$.

Acknowledgements

The authors thank Rajeev Kohli, Bari Harlam, and William Moore for their helpful comments.

References

[1] N. Anderson, Functional measurement and psychosocial judgement, *Psychol. Rev.* **77** (May 1970).
[2] B.F. Anderson, D.H. Dean, K.R. Hammond, G.H. McClelland and J.C. Shanteau, *Concept in Judgment and Decision Research* (Prager, New York, 1981).
[3] N. Anderson and J. Zalinski, A functional measurement approach to self-estimation in multiattribute environments, in: N.H. Anderson, Ed., *Information Integration Theory, Vol. I: Cognition* (L. Erlbaum, 1991) 153–170.
[4] N. Beckwith and D. Lehmann, The importance of differential weights in multiple attribute models of consumer attitude, *J. Mark. Res.* **10** (1973) 262–269.
[5] N. Beckwith and D. Lehmann, The importance of halo effects of multi-attribute models. *J. Mark. Res.* **12** (1975) 265–275.
[6] J. Bettman, *An Information Processing Theory of Consumer Choice* (Addison-Wesley, Reading, MA, 1979).
[7] J. Bettman, N. Capon, and R.L. Lutz, Multiattribute measurement models and multiattribute attitude theory: A test of construct validity, *J. Consumer Res.* **1** (1975) 1–15.
[8] R.S. Billings and S.A. Marcus, Measures of compensatory and non-compensatory models of decision behavior: Process tracing versus process capturing, *Organiz. Behav. Human Perform.* **31** (1983) 331–352.
[9] J.J. Buckley, Solving fuzzy equations in economics and finance, *J. Fuzzy Sets and Systems* **48** (1992) 289–296.
[10] R. Bucklin and S. Gupta, Brand choice, purchase incidence, and segmentation: An integrated approach, *J. Mark. Res.* **24** (1992) 201–215.
[11] T. Calinski and J. Harbasz, A dendrite method for cluster analysis, *Commun. Statist.* **3** (1974) 1–27.
[12] I.S. Currim and L.G. Schneider, A taxonomy of consumer purchase strategies in a promotion intensive environment, *Mark. Sci.* **10** (1991) 91–110.
[13] D.J. Curry and D.J. Faulds, Indexing product quality: Issues, theories, and results, *J. Consumer Res.* **13** (1986) 134–145.
[14] R.M. Dawes and B. Corrigan, Social selection based on multi-dimensional criteria, *J. Abnorm. Soc. Psychol.* **68** (1974) 104–109.
[15] G. Debreu, Representation of a preference ordering by a numerical function, in : R.M. Thrall, C.H. Coombs and R.L. Davis, Eds., *Decision Processes* (Wiley, New York, 1954).
[16] H.J. Einhorn, The use of non-linear, noncompensatory models in decision making, *Psychol. Bull.* **73** (3) (1970) 221–230.
[17] T. Elrod, Paper Presented at the 1993 Marketing Science Conference, St. Louis, 1993.
[18] P.H. Farquhar and V.R. Rao, A balance model for evaluating subsets of multiattributed items, *Management Sci.* **22** (1976) 528–539.
[19] M. Fishbein, A behavior theory approach to the relations between beliefs about an object and the attitude toward the object, in: M. Fishbein, ed., *Readings in Attitude Theory and Measurement* (John Wiley and Sons, New York, 1967).
[20] D.H. Gensch and J. Svestka, A maximum likelihood disaggregate hierarchical model for predicting choices of individuals, *J. Math. Psychol.* **28** (1984) 160–178.
[21] J.A. Goguen, L-fuzzy sets, *J. Math. Anal. Appl.* **18** (1967) 145–174.
[22] P.E. Green, *Analyzing Multivariate Data* (Dryden Press, Hinsdale, IL, 1978).
[23] P.E. Green and V. Srinivasan, Conjoint analysis in marketing: New developments with implications for research and marketing, *J. Mark.* **54** (1990) 3–19.
[24] Guadagni and J.D.C. Little, A logit model of brand choice calibrated on scanner data, *Mark. Sci.* **3** (1983) 203–238.
[25] J. Johnston *Econometric Methods* (McGraw-Hill, New York, 1984) 182–185.
[26] E.J. Johnson, R.J. Meyer and S. Ghose, When choice models fail: Compensatory models in negatively correlated environments, *J. Mark. Res.* **26** (1989) 255–270.
[27] W. Kamakura and B. Russell, A probabilistic choice model for segmentation and elasticity structure, *J. Mark. Res.* **26** (1989) 379–390.
[28] C. King and J. Summers, Overlap of opinion leadership across consumer product categories, *J. Mark. Res.* **7** (1970) 43–50.

[29] P. Kopalle and D. Hoffman, Generalizing the sensitivity conditions in an overall index of product quality, *J. Consumer Res.* **18** (1992) 530–535.

[30] K. Lancaster, *Consumer Demand: A New Approach,* (Columbia University Press, New York, 1971).

[31] D.R. Lehmann and W. Moore, A combined simply scalable and tree based preference model, *J. Bus. Res.* **22** (1991) 311–326.

[32] J.G. Lynch, Looking for confirming evidence: The case of the elusive conjunctive consumer decision process, Working Paper, University of Florida (1981).

[33] D. McFadden, Econometric models of probabilistic choice, in: C. Manski and D. McFadden, Eds., *Structural Analysis of Discrete data* (MIT Press, Cambridge, 1981) 198–272.

[34] W. Moore and D.R. Lehmann, A paired comparison nested logit model of individual preference structures, *J. Mark. Res.* **26** (1989) 420–428.

[35] B. Ratchford, The new economic theory of consumer behavior, *J. Consumer Res.* **2** (1975) 66–75.

[36] M.J. Rosenberg, Cognitive structure and attitudinal affect, *J. Abnorm. Soc. Psychol.* **53** (1956) 367–372.

[37] S.M. Shugan, The cost of thinking, *J. Consumer Res.* **7** (1981) 99–111.

[38] M. Smithson, *Fuzzy Set Analysis For the Social Sciences,* (Springer, New York, 1986)

[39] M. Smithson, Fuzzy set theory and the social sciences: The scope for applications, *Fuzzy Sets and Systems* **26** (1988) 1–21.

[40] V. Srinivasan, A conjunctive–compensatory approach to the self explication of multiattributed preferences, *Decision Sci.* **19** (1988) 295–305.

[41] V. Thole, H.J. Zimmermann and P. Zysno, On the suitability of minimum and product operators for the intersection of fuzzy sets, *Fuzzy Sets and Systems* **2** (1979) 167–180.

[42] A. Tversky, Elimination by aspects, *Psychol. Rev.* **79** (1972) 281–299.

[43] A. Tversky and S. Sattath, Preference trees, **86** (1979) 542–593.

[44] P.K. Wright, Consumer choice strategies: Simplifying vs. optimizing, *J Mark. Res.* **12** (1975) 60–67.

[45] L.A. Zadeh, Fuzzy sets, *Information and Control* **8** (1965) 338–353.

[46] J.L. Zaichkowsky, Measuring the involvement construct, *J. Consumer Res.* **12** (1985) 341–353.

[47] H.J. Zimmermann, *Fuzzy Set Theory and Its Applications* (Klewer, Leiden, 1985)

[48] H.J. Zimmermann and P. Zysno, Latent connectives in human decision making, *Fuzzy Sets and Systems* **4** (1980) 37–51.