KAMEL JEDIDI and RAJEEV KOHLI*

The authors propose two generalizations of conjunctive and disjunctive screening rules. First, they relax the requirement that an acceptable alternative must be satisfactory on one criterion (disjunctive) or on all criteria (conjunctive). Second, they relax the assumption that consumers make deterministic judgments when evaluating alternatives. They combine the two generalizations into a probabilistic subset-conjunctive rule, which allows consumers to use any number or subset of decision criteria when screening alternatives and permits them to be uncertain about the acceptability of attribute levels. These two features allow for a screening process that is uncertain and more flexible than the deterministic conjunctive and disjunctive rules currently described in the literature. The authors describe a latent-class method for the estimation of the subset-conjunctive rules and the attribute-level consideration probabilities using either consideration or choice data. Applications using both types of data suggest that the proposed models predict as well as linear models do; can make different predictions of consideration, choice, and market shares; and provide insights into consumer decision processes that are different from those obtained with linear models.

Probabilistic Subset-Conjunctive Models for Heterogeneous Consumers

> [T]here may be some behaviors for which the primitive attributes do not compensate but act in a manner I shall call conjunctive behavior. By this I mean behavior in which several attributes are required, each to a minimum degree…. This type of behavior has an obvious counterpart in the mathematical model of what can be called disjunctive behavior, in which the primitive attributes are combined on an "either–or" basis.
>
> —Clyde Coombs (1951)

The use of conjunctive and disjunctive rules that Coombs (1951) describes in the opening quotation is now well documented in consumer research. Protocol analyses and self-reports suggest that consumers use these rules to screen alternatives, for example, when they evaluate large choice sets or make complex or unfamiliar decisions (for a review, see Bettman 1979). Such a screening separates the alternatives into "acceptable" and "unacceptable" categories. Einhorn (1970), Abe (1999), Swait (2001), and Elrod, Johnson, and White (2004) describe nonlinear representations of conjunctive and disjunctive rules. The substantial literature on choice set formation suggests that consumers do not choose from a universal set of alternatives but rather from a reduced consideration set (see Roberts and Nedungadi 1995). The works of Gensch (1987) and Gilbride and Allenby (2004) are examples of two-stage models in which conjunctive or disjunctive screening precedes choice.

This article proposes (1) a generalization of conjunctive and disjunctive rules, (2) an approach to accommodating uncertainty in consumers' use of these more general rules, and (3) a procedure for inferring segment-level rules that reflect the preceding two generalizations. We restrict our models to discrete, possibly ordered attributes. We now briefly discuss the three aspects of the proposed rules and the method for their inference.

Generalization of Rules

We describe a rule by which an acceptable alternative must satisfy at least t of m possible decision criteria. The value t = 1 corresponds to a disjunctive rule, and the value t = m corresponds to a conjunctive rule. As the value of t increases from 1 to m, we obtain rules that require the satisfaction of increasingly more criteria than does a disjunctive rule. We call this generalization a "subset-conjunctive rule." A special case of the rule occurs when a consumer uses a conjunctive rule over a fixed subset of t out of m possible decision criteria. This type of rule might occur, for example,

*Kamel Jedidi is Professor of Marketing (e-mail: kj7@columbia.edu), and Rajeev Kohli is Professor of Marketing (e-mail: rk35@colombia.edu), Graduate School of Business, Columbia University. The authors thank the three anonymous *JMR* reviewers for their constructive suggestions and Sawtooth Software, McKinsey & Company, and Intelliquest for providing the personal computer choice data. They also thank Greg Allenby and Tim Gilbride for helpful discussions and clarifications regarding their work.

ISSN: 0022-2437 (print), 1547-7193 (electronic)

when a consumer is shown alternatives with m attributes but uses only a subset of the attributes in his or her evaluation.

A feature of subset-conjunctive rules, excluding the conjunctive rule, is that they can be used for evaluating alternatives for which the available information is incomplete on the decision criteria. For example, suppose a consumer requires that an acceptable alternative is satisfactory for $t = 2$ attributes. He or she can then make an acceptable classification if information on only two attributes is available, as long as the alternative is acceptable on both attributes; otherwise, the consumer classifies the alternative as unacceptable. For this reason, subset-conjunctive rules are a variety of the more general class of partially defined Boolean functions (Crama, Hammer, and Ibaraki 1988). Subset-conjunctive rules also define a subclass of models in the logical analysis of data (Boros, Hammer, and Hooker 1995), which is used in discrimination problems outside consumer decision making (e.g., in the diagnosis of diseases; see, e.g., Alexe et al. 2003).

Consumer Uncertainty

All existing models of which we are aware assume deterministic conjunctive/disjunctive screening rules. They do not allow for uncertainty in judgments or error in the use of the rules; therefore, they imply that only a single consideration set needs to be evaluated in choice models that incorporate these screening rules. We introduce uncertainty in conjunctive/disjunctive rules by associating a "consideration probability" with which a consumer finds an attribute level to be acceptable. A consumer associates a greater consideration probability with an attribute level that he or she is more likely to find acceptable. These probabilities can reflect uncertainties in tastes or the effect of lack of information or knowledge. An example of the first is a consumer who is unsure about whether to choose a Japanese or an Indian restaurant for lunch. An example of the second type is a consumer who does not know enough about plasma screen television and receives conflicting recommendations from experts. We allow the consideration probabilities to differ across attribute levels: A \$5,000 price for a new stove may be unacceptable, a \$2,000 price may be acceptable, and intermediate prices may be acceptable to varying degrees.

The proposed level-by-level variation in consideration probabilities is a different form of uncertainty than is reflected by a standard formulation of an additive error term in a utility function. It is akin to the probabilities by which consumers select aspects in Tversky's (1972) elimination-by-aspects model, except that we do not assume a relationship between a consideration probability and the frequency with which an attribute level appears across alternatives.

Segmentation

We propose a latent-class model for inferring probabilistic subset-conjunctive rules for unobserved consumer segments. The rules that we obtain reflect a combination of consumer uncertainty, response error, and within-segment heterogeneity. An advantage of assessing segment-level rules is that they provide a summary of similarity in the types of rules that consumers use. Individual-level rules can become too numerous, and it can become difficult to observe patterns across these rules. Conversely, individual-level rules, such as those that Gilbride and Allenby (2004) infer, can be more useful in conjoint simulations and for customizing offerings to consumers. The two types of models offer complementary analyses.

We close this section by noting that linear models are sufficient for representing deterministic, error-free conjunctive/disjunctive types of rules. However, as we show in the Appendix, the parameters of these models are not unique. This suggests that if there is low response error, it may not even be possible to estimate linear models with data obtained from conjunctive/disjunctive processes. Estimation becomes possible as response error increases, and it may be possible to obtain good fit with linear models. However, as we discuss in the "Application of the Choice Model" section, this does not necessarily imply that a linear model will make the same predictions and recommendations as a conjunctive/disjunctive type of model.

Organization of the Article

We begin the next section by describing the proposed subset-conjunctive rule and its probabilistic generalization. We then describe a method for inferring unobserved consumer segments and the associated consideration rules using data on acceptable/unacceptable classification of multiattribute alternatives. Next, we present an application of the method using consideration acceptable/unacceptable data for household batteries. After this, we introduce the two-stage choice model. In this model, consideration and choice are two separate but related steps. We represent consideration by an unknown probabilistic subset-conjunctive rule. This rule implies a probability of consideration with each possible subset of items in a choice set. The second stage is the selection of an alternative from a consideration set. We model this stage using a logit formulation in which the first-step consideration probabilities are the predictor variables. We then illustrate the choice model and compare it with a Dogit model using data from a personal computer (PC) choice experiment.

PROBABILISTIC SUBSET CONJUNCTION

Consider a single consumer who classifies N alternatives into acceptable/unacceptable categories. We assume that each alternative is defined over $m(\geq 2)$ attributes. Let attribute k have $n_k(\geq 2)$ levels, $1 \leq k \leq m$. The consumer considers each attribute level satisfactory or unsatisfactory. Consider that the consumer uses a t-subset-conjunctive rule if each acceptable alternative is satisfactory on at least t of the m attributes, $1 \leq t \leq m$; otherwise, the alternative is unacceptable. The value of t is fixed for the consumer, but it can vary across consumers. Different values of t define rules that differ in the "hurdle" that the consumer imposes on acceptable alternatives. The special cases $t = 1$ and $t = m$ correspond to the disjunctive and conjunctive rules. Another special case is a conjunction over a given subset of t attributes. The consumer then either does not consider the remaining $m - t$ attributes or finds each of their levels acceptable.

We now consider a probabilistic generalization of a subset-conjunctive rule. We define η_{jk} as the probability that a consumer finds level j of attribute k acceptable. We call η_{jk} the consideration probabilities. We assume that the consideration probabilities are pairwise independent. This is a probabilistic version of the standard assumption (e.g., Gilbride and Allenby 2004) that consumers independently evaluate the attribute levels when using a conjunctive/

disjunctive rule. For ordered attributes, we allow non-decreasing consideration probabilities when the levels are arrayed in increasing preference order. The special case in which the η_{jk} have 0 – 1 values reduces to a deterministic subset-conjunctive rule. Deviations from these extreme values represent uncertainties in judgment about the acceptability of an attribute level.

The values of η_{jk} provide information about the importance of an attribute in product screening by a consumer. Attribute k is irrelevant if $\eta_{jk} = 0$ or if $\eta_{jk} = 1$ for all levels $1 \le j \le n_k$. If $\eta_{jk} = \eta$ for all $1 \le j \le n_k$, where η is the proportion of alternatives acceptable to a consumer, then attribute k has no discriminating ability. These properties are useful for interpreting the parameters of the proposed model.

Next, we specify the relationship between the consideration probabilities and the acceptable/unacceptable classifications of alternatives. Let

$$(1) \qquad x_{ijk} = \begin{cases} 1 & \text{if alternative i has level j of attribute k,} \\ 0 & \text{otherwise;} \end{cases}$$

and let

$$(2) \qquad y_i = \begin{cases} 1 & \text{if a consumer finds alternative i acceptable,} \\ 0 & \text{otherwise.} \end{cases}$$

Let p_{ik} ($q_{ik} = 1 - p_{ik}$) denote the probability that a consumer finds alternative i acceptable (not acceptable) on attribute k. Because each alternative has one level of each attribute, we can write p_{ik} as

$$(3) \qquad p_{ik} = \prod_{j=1}^{n_k} \eta_{jk}^{x_{ijk}} \text{ for all } 1 \le k \le m.$$

Let π_{it} denote the probability that a consumer using a t-subset-conjunctive rule finds alternative i acceptable. Then, $\pi_{it} = f(\eta_{jk})$ is a function of the consideration probabilities. Suppose we array the m attributes in increasing order of their (arbitrary) indices. Then, there are k – 1 attributes preceding attribute k. We consider the m – (k – 1) attributes with indexes no smaller than k. Let $\theta_{is}(k, m)$ denote the probability that at least s of these m − k + 1 attributes are acceptable to the consumer. If k = 1 and s = t, then $\theta_{it}(1, m)$ denotes the probability that the consumer finds alternative i acceptable on at least t of attributes 1 to m; that is,

$$(4) \qquad \pi_{it} \equiv \theta_{it}(1, m).$$

Next, we write an expression for $\theta_{it}(1, m)$ by considering the following two possibilities: First, we consider the outcome, which has probability p_{i1}, that alternative i is acceptable on attribute 1. Then, alternative i is acceptable if it is also acceptable on t – 1 or more of attributes 2 to m; the probability that this occurs is $\theta_{i(t-1)}(2, m)$. Second, we consider the outcome, which has probability q_{il}, that alternative i is not acceptable on attribute 1. Then, alternative i is acceptable only if it is acceptable on t or more of attributes 2 to m; the probability that this occurs is $\theta_{it}(2, m)$. Combining the two possibilities, we obtain

$$(5) \qquad \pi_{it} = \theta_{it}(1, m) = p_{i1}\theta_{i(t-1)}(2, m) + q_{i1}\theta_{it}(2, m).$$

We now expand each θ term on the right-hand side in a recursive fashion, using the more general expression that is analogous to Equation 5:

$$(6) \qquad \theta_{is}(k, m) = p_{ik}\theta_{i(s-1)}(k+1, m) + q_{ik}\theta_{is}(k+1, m), \text{ for all } 1 \le k \le m,$$

where, by definition,

$$(7) \qquad \theta_{is}(k, m) \equiv 0, \text{ for all } s > m - k + 1.$$

We continue to use the recursion until the final expression contains only the p_{ik} terms, at which point we substitute for the η_{jk} using Equation 3.

Rule Inference

Suppose we know the acceptable/unacceptable classifications of the N alternatives by a consumer. If we assume a value of t, we can use the preceding method to write the probability π_{it} that alternative i is acceptable to the consumer. Thus, we can write the expression for the joint probability that the consumer finds a certain subset of alternatives acceptable and the other subsets unacceptable. We can write this joint probability (i.e., the likelihood for the consumer) as follows:

$$(8) \qquad L_t = \prod_{i=1}^{N} \pi_{it}(1 - \pi_{it}^{y_i})^{1 - y_i}.$$

The right-hand side of Equation 8 is an implicit function of the consideration probabilities η_{jk}. We maximize Equation 8 with respect to the η_{jk}, which is subject to the constraint that for ordered attributes, the value of η_{jk} is no smaller for a preferred level. We now repeat the analysis for different values of t between 1 and m. We select the value of t for which L_t has the highest value.

Segment-Level Extension

We now consider the extension to G unobserved segments. The segment-level model is identified only if each consumer h evaluates $N_h \ge 2$ alternatives.

We modify the expressions in Equations 3 to 8 in an obvious manner, adding a superscript g to refer to segment-level terms. We write the unconditional likelihood function for consumer h as

$$(9) \qquad L_{ht}^G = \sum_{g=1}^{G} w_g \times L_{ht}^g,\ w_g \ge 0,\ \sum_{g=1}^{G} w_g = 1,$$

where, analogous to Equation 8, L_{ht}^g is the likelihood function conditional on membership in segment g and w_g is the prior probability (mixing proportion) that consumer h belongs to latent segment g. The likelihood function across H consumers is then given by

$$(10) \qquad L_t^G = \prod_{h=1}^{H} L_{ht}^G.$$

We maximize Equation 10 with respect to w_g and η_{jk}^g, $1 \le g \le G$, for given values of G and t.

Issues in Model Estimation and Selection

We use proc NLP in SAS to maximize Equation 10. This optimization procedure allows for linear constraints on the η_{jk} for ordered attributes, permits boundary conditions, and provides standard errors for the consideration probabilities and mixing proportions. Although Equations 8 and 10 can be high-order polynomials, we encountered no convergence problems when using the method for maximizing the likelihood functions.[1] However, it is advisable to check for the possibility of local optima by using several different starting values for the parameters.

The number of segments, G, and the subset size, t, are unknown. Therefore, we select some reasonably high maximum value for G (e.g., we evaluate six segments in the application we report next) and estimate the finite mixture model for all G × m combinations. We select the solution that minimizes the Bayesian information criterion (BIC):

$$\text{(11)} \qquad BIC_t^G = -2\log L_t^G + M_G \log\left(\sum_{h=1}^{H} N_h\right),$$

$$M_G = (G-1) + G\sum_{k=1}^{m} n_k,$$

where M_G is the number of effective parameters estimates. We use Bayes' rule to compute the probability that consumer h belongs to segment g:

$$\text{(12)} \qquad P_{hg} = \frac{w_g L_{ht}^g}{\sum_{l=1}^{G} w_l L_{ht}^l}.$$

We evaluate the separation of the segments using an entropy measure (Wedel and Kamakura 1999).

$$\text{(13)} \qquad E_G = 1 - \frac{1}{H\log G}\sum_{h,g} P_{hg}\log P_{hg},\ 0 \le E_G \le 1.$$

A larger entropy value indicates better separation.

Table 1 summarizes the main differences between our proposed model and the threshold-based approach that Gilbride and Allenby (2004) use. We estimate probabilistic subset-conjunctive rules for segments of consumers. In contrast, the first stage of Gilbride and Allenby's (2004, p. 393, Eqs. 3 and 4) model is restricted to deterministic conjunctive or disjunctive rules that can vary across individual consumers. The distribution of rules is obtained by allowing a multinomial distribution for attribute-specific thresholds (see Gilbride and Allenby 2004, p. 398, Eq. 21, and their Appendix). Only attribute levels passing a threshold are acceptable for a person. This formulation implies ordinal preferences over attributes; for example, it is not possible to sequence three levels of an attribute such that a single threshold makes the first and third levels acceptable but not the second level. In contrast, our approach allows for both ordinal and nominal attributes. The second (choice) stage in Gilbride and Allenby's model is invoked only when one or more alternatives pass the screening step. A consumer chooses an alternative with the greatest multiattribute utility if it meets two conditions: (1) It is not an item already eliminated at the first stage, and (2) it has a greater utility value than a "no-choice" option.[2] Because it is a multinomial choice model, Gilbride and Allenby's two-step approach

[1]Simulation results suggest the accuracy of the estimation procedure for recovering the "true" model parameters. Across 20 experimental conditions, in which we varied the amount of error (2 levels), sample size (2 levels), and subset size (5 levels) and ran 100 replications per cell, the mean absolute deviation between the true and the estimated parameters is .04 and the percentage of trials indicating the true subset size t is 93.2.

[2]Thus, no choice can occur if all alternatives are rejected in the first stage or if each alternative evaluated in the second stage has a lower utility than a no-choice option.

Table 1

COMPARISON OF PROPOSED MODEL OF CONSIDERATION AND GILBRIDE AND ALLENBY'S (2004) MODEL OF CONSIDERATION AND CHOICE

Criteria	*Gilbride and Allenby's Model*	*Proposed Model*
Type of consideration rule	•Conjunctive or disjunctive rule. •Deterministic, person-specific thresholds.	•Probabilistic subset-conjunctive rule. Associates a consideration probability with each attribute level. An acceptable alternative must be satisfactory on at least t of m criteria, 1 ≤ t ≤ m. •Subsumes deterministic conjunctive and disjunctive rules as special cases.
Dependent variable	•Choice of at most one item from a set of alternatives. •Assumes independence between consideration and choice functions.	•Binary (acceptable/unacceptable) classifications of alternatives. •Allows a two-stage choice model in which choice probabilities and consideration probabilities are not independent.
Independent variables	•Ordinal attributes.	•Nominal and/or ordinal attributes.
Level of analysis	•Infers individual-level rules from threshold estimates that determine the acceptability of each attribute level.	•Segment-level parameter estimates. Useful for assessing group-level similarities in consideration rules.
Estimation	•No error model for consideration stage. •Reflects heterogeneity in consideration rules by assuming a multinomial distribution of thresholds across individuals. •Person-specific estimates may be unreliable if there are not enough observations per individual.	•Parameter estimates reflect uncertainty in consideration and choice. •Segment-level estimation may allow fewer observations per person.

appears to subsume the proposed binary (acceptable/unacceptable) model as a special case for the deterministic conjunctive and disjunctive rules. However, the use of two stages for classifying a single item into acceptable/unacceptable categories seems somewhat forced and redundant (and is likely not intended by the authors). In the "Extension to Choice" section, we describe an extension of our proposed consideration model to a two-step choice model.

APPLICATION OF CONSIDERATION MODEL

A manufacturer of household batteries conducted a study to assess consumer interest in possible battery improvements. A total of 175 consumers, who were screened to be battery purchasers, were interviewed at shopping malls. Each participant was shown in random order descriptions of 32 battery concepts that were written on cards. The respondents' task was to examine each concept one at a time and indicate whether they would consider purchasing the battery.[3] The 32 battery profiles were generated from an experimental plan that confounded the main effects and (second-order and higher) interaction effects for six design factors: (1) incremental price (0%, 25%, and 50% higher than the current price), (2) built-in charge meter (yes/no), (3) environmental safety (yes/no), (4) rapid recharge (yes/no), (5) life of the battery (standard, 50% longer than the standard), and (6) brand name (which we call A, B, C, and D). On average, respondents did not consider 18% of the battery profiles.

Let x_{21} and x_{31} denote dummy variables representing the 25% and 50% higher price levels. Let x_{22}–x_{25} be dummy variables representing built-in charge meter, environmental safety, rapid recharge, and battery life, respectively (see Table 2). The "brand" attribute is not statistically significant ($p \geq .05$) for all subset-conjunctive and logistic regression models; thus, we do not refer to it in our subsequent discussion.

Logistic Regression

A logistic regression of the pooled data has a log-likelihood (LL) value of –2003.8 and estimates the following utility function:

$$u = 1.7 - 1.57x_{21} - 3.1x_{31} + .88x_{22} + 1.07x_{23} + .75x_{24} + 1.21x_{25}.$$

All parameter estimates are significant at $p < .05$. These results suggest that price is the most important factor, though its relative importance may be inflated because it has the most levels among the attributes (Verlegh, Schifferstein, and Wittink 2002). Longer battery life is the most important of the remaining factors, followed by environmental safety, the importance of which might also be inflated because we use hypothetical product concepts. These results do not appear to suggest a subset-conjunctive rule.

Next, we consider the results from a latent-class logistic regression. The BIC suggests a three-segment solution

[3]The simplicity of the consider/do not consider task mitigates, at least in part, the possibility of a respondent's artificial processing of 32 product profiles.

Table 2

POOLED (G = 1) PARAMETER ESTIMATES FOR THE SUBSET-CONJUNCTIVE CONSIDERATION MODEL IN THE BATTERY STUDY

		Label		*Subset Size*				
Attribute	*Level*	*Variable*	*Parameter*	*1*	*2*	*3*	*4*	*5*
Price	0%	x_{11}	η_{11}	.88 (.02)	1.00 (bc)	1.00 (bc)	1.00 (bc)	1.00 (bc)
	25%	x_{21}	η_{21}	.55 (.02)	.69 (.03)	.73 (.03)	.76 (.03)	.97 (.01)
	50%	x_{31}	η_{31}	.00 (bc)	.00 (bc)	.00 (bc)	.00 (bc)	.70 (.02)
Built-in meter	Yes	x_{22}	η_{22}	.40 (.04)	.53 (.03)	1.00 (bc)	1.00 (bc)	1.00 (bc)
	No	x_{12}	η_{12}	.00 (bc)	.00 (bc)	.70 (.03)	.79 (.02)	.94 (.01)
Environmental safety	Yes	x_{23}	η_{23}	.49 (.03)	.66 (.03)	.95 (.03)	1.00 (bc)	1.00 (bc)
	No	x_{13}	η_{13}	.0 (bc)	.14 (.03)	.55 (.04)	.75 (.02)	.94 (.01)
Rapid recharge	Yes	x_{24}	η_{24}	.38 (.04)	.94 (.02)	.95 (.02)	1.00 (bc)	1.00 (bc)
	No	x_{14}	η_{14}	.07 (.02)	.76 (.03)	.68 (.04)	.81 (.02)	.96 (.01)
Longer life	Yes	x_{25}	η_{25}	.54 (.03)	.62 (.03)	.60 (.03)	.98 (.01)	1.00 (bc)
	No	x_{15}	η_{15}	.00 (bc)	.00 (bc)	.00 (bc)	.77 (.03)	.94 (.01)
LL				–2011	–1999[a]	–2023	–2078	–2215

[a]Denotes best pooled (G = 1) solution based on maximum LL value.
Notes: Standard errors are in parentheses; bc = standard error cannot be computed because of the boundary condition.

(LL = –1454.1). The segment-level utility function u_g for segment g is

$$u_1 = .75 - 1.28x_{21} - 2.79x_{31} + 1.93x_{22} + 2.35x_{23} + 1.12x_{24} + 3.38x_{25},$$

$$u_2 = 1.31 - 6.57x_{21} - 6.80x_{31} + 1.25x_{22}, \text{ and}$$

$$u_3 = 2.92 - 1.77x_{21} - 4.73x_{31} + .55x_{22} + .79x_{23} + .63x_{24} + .45x_{25}.$$

Each of the parameter estimates in the preceding expressions is significant at the $p = .05$ level. The large negative parameter estimate for x_{21} in Segment 2 implies that for any battery, the consideration probability drops to nearly zero at a 25% higher price. A similar observation applies to the 50% higher price for Segments 2 and 3. For these prices, the logistic regression itself suggests a noncompensatory evaluation process, but for other prices, the model suggests varying trade-offs across segments.

Probabilistic Subset Conjunction

The last row of Table 2 gives the LL values for the probabilistic subset-conjunctive models with pooled data. The highest LL value is for the t = 2 solution, and this value is slightly higher than the LL value for the pooled logistic regression. There is substantial within-attribute variation in the $\hat{\eta}_{jk}$ values. The least preferred attribute levels are unacceptable, or nearly so. This solution suggests that an alternative needs to be acceptable on at least two attributes for a consumer to consider it. For example, using Equation 5, we obtain a probability of .95 that a consumer will consider a battery with a regular price and rapid recharge but with no built-in meter, no environmental safety, and no longer life.

We now examine the results for the latent-class subset-conjunctive model. The BIC suggests a three-segment solution with subset size t = 2. Both the logistic regression and the subset-conjunctive model produce almost identical segment memberships and equal entropies of .93, suggesting excellent separations among the segments. More than 60% of the respondents are in Segment 1, more than 33% are in Segment 3, and the rest are in Segment 2. Table 3 shows the t = 1 and t = 2 solutions for each of the three segments. In all cases, we find that a preferred level of an attribute has a greater value for the estimated consideration probability. The t = 2 solution has a greater LL, and therefore it seems to be the better solution. However, the disjunctive solution appears to be more appropriate for Segment 3 because when t = 2, the estimated consideration probabilities are quite close for the "yes" ($\hat{\eta}_{24}$ = .98) and "no" ($\hat{\eta}_{14}$ = .91) levels of the rapid recharge attribute.

Table 3

THREE-SEGMENT PARAMETER ESTIMATES FOR THE SUBSET-CONJUNCTIVE CONSIDERATION MODEL IN BATTERY STUDY

			Segment					
			Disjunctive (t = 1)			*Subset-Conjunctive (t = 2)*		
Attribute	*Level*	*Parameter Label*	*1*	*2*	*3*	*1*	*2*	*3*
Price	0%	η_{11}	.72 (.04)	.93 (.04)	.97 (.01)	.87 (.05)	1.00 (bc)	1.00 (bc)
	25%	η_{21}	.38 (.04)	.00 (bc)	.86 (.02)	.50 (.06)	.00 (bc)	.89 (.02)
	50%	η_{31}	.00 (bc)	.00 (bc)	.12 (.07)	.07 (.03)	.00 (bc)	.00 (bc)
Built-in meter	Yes	η_{22}	.73 (.04)	.03 (.02)	.11 (.07)	.83 (.03)	.03 (.03)	.16 (.06)
	No	η_{12}	.00 (bc)	.00 (bc)	.00 (bc)	.00 (bc)	.00 (bc)	.00 (bc)
Environmental safety	Yes	η_{23}	.80 (.10)	.01 (bc)	.29 (.06)	.89 (.03)	.01 (.03)	.44 (.05)
	No	η_{13}	.00 (bc)	.00 (bc)	.02 (.06)	.00 (bc)	.00 (bc)	.18 (.05)
Rapid recharge	Yes	η_{24}	.41 (.02)	.03 (.02)	.16 (.04)	.95 (.03)	.03 (.03)	.98 (.01)
	No	η_{14}	.00 (bc)	.00 (bc)	.00 (bc)	.82 (.04)	.00 (bc)	.91 (.03)
Longer life	Yes	η_{25}	.89 (.02)	.00 (bc)	.16 (.06)	1.00 (bc)	.97 (.03)	.16 (.07)
	No	η_{15}	.00 (bc)	.00 (bc)	.00 (bc)	.00 (bc)	.89 (.05)	.00 (bc)
Mixing proportions		w_g	.61	.05	.34	.62	.05	.33
Number of respondents			107	9	59	109	9	57
LL				–1483			–1476	

Notes: Standard errors are in parentheses; bc = standard error cannot be computed because of the boundary condition.

We now briefly discuss the t = 2 solution for Segments 1 and 2 and the t = 1 solution for Segment 3. Consumers in Segment 1 appear to be the most responsive to the inclusion of new features. For example, a battery with the most desirable levels on all attributes has a consideration probability of .97, even if it has a 50% higher price than current batteries.[4] Longer life alone raises the probability of consideration to at least .9, and environmental safety raises the consideration probability to at least .8 for consumers in this segment. In contrast, only price affects the probability of consideration for a battery offered to consumers in Segment 2. This probability is nearly one for batteries offered at regular price. This suggests that consumers in Segment 2 consider only price when assessing the acceptability of alternatives. Consumers in Segment 3 are similar to those in Segment 2, except that (1) they also consider batteries sold at a 25% higher price with a reasonably high probability (≥.86), and (2) they are more responsive to the inclusion of new features. At 50% higher prices, their consideration probability is as low as 14% with all base-level features (i.e., 1 – [{1 – .12}{1 – 0}{1 – .02}{1 – 0}{1 – 0}] = .14); it rises to .61 for a battery with environmental safety, longer life, and rapid recharge.

We use the following method to compare the predictive accuracies of the three-segment logistic regression solution with the t = 1 and t = 2 subset-conjunctive solutions. We randomly select 26 product profiles from each consumer for model estimation and use the remaining 6 profiles for prediction. We use the appropriate model to compute the probability that a holdout profile is acceptable and use a probability cutoff of .5 to predict consideration. The mean hit rates across randomly drawn holdout profiles are .89, .90, and .90 for logistic regression, subset two, and disjunctive models, respectively. All three values exceed the 82% hit rate predicted by the maximum chance criterion. For each of the three models, we also compare the observed proportions of the 32 profiles that were considered in the holdout sample with the corresponding values of the average predicted consideration probability. All three models perform equally well in recovering the actual consideration proportions; the mean absolute deviation is equal to .04. These results confirm our expectation that a linear model is adequate for representing subset-conjunctive rules (see the Appendix). As we noted previously, the value of the proposed probabilistic subset-conjunctive model does not come from improved prediction but rather from the insight it provides about a consumer decision process.

EXTENSION TO CHOICE

Consider a consumer who uses a deterministic t-subset-conjunctive rule for product screening. It is possible that one considered alternative is acceptable on exactly t attributes and that another considered alternative is acceptable on all m attributes. The latter alternative may have a greater probability of subsequent choice, but there is no way to determine this from the consideration rule itself or from the results of its use for screening alternatives. In contrast, the proposed probabilistic subset-conjunctive rule allows for the possibility that choice probabilities may be related to consideration probabilities. For example, the estimates in Table 3 imply that the probability of consideration for a battery increases when it has more desirable features and a lower price. We expect the same directional effects on choice probabilities. Thus, it is possible that consideration probabilities contain information about choice probabilities. We describe such a model subsequently. Gilbride and Allenby (2004) describe an alternative choice model in which consideration and choice are independent stages. It is relatively straightforward to replace their deterministic conjunctive/disjunctive rules with the proposed probabilistic subset-conjunctive rules. Therefore, we do not examine such a formulation further in this article.

Let S denote a choice set with $|S| = R$ alternatives. If a consumer makes no choice, we assume that he or she eliminates all items in the consideration stage.[5] Assuming that each alternative is independently evaluated for consideration, the no-choice probability for a consumer using a t-subset-conjunctive rule is given by

(14) $$P_t(\bar{S}) = \prod_{i=1}^{R}(1 - \pi_{it}),$$

where, from Equation 5, π_{it} denotes the consideration probability for the alternative $i \in S$. Let ω denote a nonempty subset of alternatives in S, and let Ω denote the set of all such (nonempty) subsets. The probability that a consumer considers only the items in $\omega \in \Omega$ is given by

(15) $$\pi_t(\omega) = \prod_{i \in \omega}\pi_{it}\prod_{i \in S\backslash\omega}(1 - \pi_{it}).$$

Let $P_{it}(\omega)$ denote the probability that alternative $i \in \omega$ is selected from the subset ω. We subsequently consider the relationship between $P_{it}(\omega)$ and the consideration probability π_{it}. Here, we note that when $P_{it}(\omega)$ is specified, we can write the unconditional choice probability for alternative i as follows:

(16) $$P_{it}(S) = \sum_{\omega \in \Omega} P_{it}(\omega) \times \pi_t(\omega), \ \ P_{it}(\omega) \equiv 0 \text{ if } i \notin \omega.$$

Let $y_i = 1(0)$ if the consumer chooses (does not choose) alternative $i \in S$. We can write the likelihood of selection for alternative $i \in S$ as follows:

(17) $$L_t(S) = P_t(\bar{S})^{1 - \Sigma_i y_i}\prod_{i \in S}P_{it}(S)^{y_i}.$$

Let Ψ denote the collection of (possibly repeated) choice sets evaluated by a consumer. The likelihood function for the consumer is

(18) $$L_t = \prod_{S \in \Psi} L_t(S).$$

The right-hand side of Equation 18 is an implicit function of the consideration probabilities η_{jk}. We can estimate η_{jk}

[4]We compute all considerations probabilities using Equation 5, given the η_{jk} and t-values from Table 2 for the segment.

[5]This contrasts with the assumption by Gilbride and Allenby (2004) that there are two possible reasons for no choice: rejection of all items at the consideration stage, and rejection of all considered items at the choice stage.

by maximizing Equation 18 for each $1 \leq t \leq m$ and by selecting the solution for which L_t is the largest. The parameters η_{jk} are uniquely identified only if the data collection allows for a no-choice option for each choice set. Intuitively, without a no-choice option, it is not possible to estimate the probability that no alternative is acceptable.

Specification of Conditional Choice Probabilities

As we noted previously, we still need to specify the relationship between the conditional choice probabilities $P_{it}(\omega)$ and the consideration probabilities π_{it}. Let $U_i = V_i + \varepsilon_i$, where $V_i \equiv V_i(\pi_{it})$ and ε_i is a random error term, which we assume follows an extreme value distribution. Then, the conditional choice probability is given by

$$(19) \qquad P_{it}(\omega) = \frac{\exp(V_i)}{\sum_{\ell \in \omega} \exp(V_\ell)}, \ \omega \in \Omega.$$

We tested two possible relationships: $V_i = \beta \log \pi_i$ and $V_i = \beta\pi_i$. Empirically, the former specification does substantially better. Therefore, we use it in the application we describe in the next section.[6]

Segment-Level Extension

Let G denote the number of segments. Let consumer h evaluate $N_h \geq 2$ choice sets. Then, we can extend the preceding model to infer latent-class, subset-conjunctive rules. We modify the expressions in Equations 14–19 in an obvious manner, adding a superscript g to refer to segment-level terms and a subscript h to refer to a particular consumer. Analogous to Equation 16, the unconditional choice probability for alternative i by a consumer in segment g is given by

$$(20) \qquad P_{it}^{g}(S) = \sum_{\omega \in \Omega} P_{it}^{g}(\omega) \times \pi_{it}^{g}(\omega), \ \ P_{it}^{g}(\omega) \equiv 0 \text{ if } i \notin \omega.$$

The unconditional likelihood function for consumer h is

$$(21) \qquad L_{ht}^{G} = \sum_{g=1}^{G} w_g \times L_{ht}^{g}, \ w_g \geq 0, \ \sum_{g=1}^{G} w_g = 1,$$

where, analogous to Equation 18, L_{ht}^{g} is the likelihood function conditional on membership in segment g and w_g is the prior probability (mixing proportion) that consumer h belongs to latent segment g. The likelihood function across H consumers is then given by

$$(22) \qquad L_{t}^{G} = \prod_{h=1}^{H} L_{ht}^{G}.$$

We maximize Equation 22 using proc NLP in SAS for each of $G \times m$ combinations of the number of segments and the subset sizes. We select a solution that minimizes the BIC and obtain estimates for the subset-size t*; the number of segments G*; the mixing proportions w_g, g = 1, ..., G; and the segment-level consideration probabilities η_{jk}^{g}.

APPLICATION OF THE CHOICE MODEL

We illustrate the preceding choice model using data from an experiment on PC choices. Each of 297 participants was shown eight choice sets and was asked to choose, at most, one PC from each choice set. The choice sets differed across participants and were selected from a master experimental plan. Each item in a choice set was described over the following five attributes: (1) brand name (A, B, C, D, E), (2) performance (below average, average, above average), (3) warranty period (90 days, 1 year, 5 years), (4) service location (ship back to manufacturer, service at local dealer, on-site service), and (5) price (low, medium-low, medium-high, high). We estimate the preceding choice model and compare its results with those from a latent-class version of a nested Dogit model (Gaudry and Dagenais 1979).[7] The Dogit model subsumes the nested multinomial logit model as a special case and permits captivity when only one alternative is considered and chosen, no choice when no alternative is considered, and choice among all brands.

The BIC indicates a two-segment solution for both the nested Dogit and the subset-conjunctive choice model with t* = 4. The entropy for the two-segment solution is .7 for both models, indicating good separation of the segments. The BIC values and the LL values are substantially smaller for the proposed model (LL = –2156.7; BIC = 4616.5) than for the nested Dogit model (LL = –2197.1; BIC = 4697.4), suggesting a better fit for the proposed model. To assess the accuracy of holdout predictions, we estimate the parameters using a random 90% of the choice sets and then predicting brand choices for the 10% remaining choice sets. We repeat the analysis 100 times. The mean absolute deviation between the actual choice proportions and the predicted choice probabilities is .015 for the proposed model and is .016 for the nested Dogit model. Thus, both models appear to yield comparable holdout predictions.

Table 4 shows the pooled (G = 1) and segment-level parameter estimates for the proposed choice model. One way to assess the relative importance of an attribute is by comparing the range of consideration probabilities across their respective levels. By this measure, "performance" is the most important attribute for both segments. "Service" is the second-most important attribute for Segment 1 (range = .76), but it is much less important for Segment 2 (range = .2). Segment 2 has higher estimates of the consideration probabilities because it has a smaller proportion of no choices. The larger value of the choice exponent β for Segment 2 can be shown to skew the probabilities toward the extremes for two of the three alternatives in a choice set. It implies that compared with Segment 1, changing an attribute from the most to the least preferred level has a smaller effect on the consideration probability for Segment 2 (because of smaller ranges) but a greater effect on choice (because of a larger value of β).

[6]We do not pursue a more general examination of alternative functional relationships between $P_{it}(\omega)$ and π_{it}, which would take us too far afield. It is worth noting here that we have also empirically tested the additive specification $V_i = \mathbf{x}_i\beta$, where $\mathbf{x}_i$ is a vector of predictor variables for choice. Such a specification is not parsimonious, and we rejected it in our empirical analysis using the BIC.

[7]We are grateful to an anonymous reviewer for suggesting the Dogit model as a benchmark for comparing the proposed choice model. A working paper, which is available on request, provides the details of the Dogit model and its estimation.

Table 4

POOLED (G = 1) AND TWO-SEGMENT PARAMETER ESTIMATES FOR THE SUBSET-CONJUNCTIVE CHOICE MODEL IN PC STUDY

			Attribute-Level Consideration Probabilities for the Subset-Conjunctive Model (t = 4)					
Attribute	*Level*	*Label*	*Pooled*	*Standard Error*	*Segment 1*	*Standard Error*	*Segment 2*	*Standard Error*
Brand	A	x_{11}	.66	.08	1.00	(bc)	.55	.17
	B	x_{21}	.59	.08	.95	.10	.50	.18
	C	x_{31}	.56	.08	.91	.09	.47	.18
	D	x_{41}	.44	.08	.70	.10	.38	.19
	E	x_{51}	.35	.09	.58	.09	.30	.20
Performance	Below average	x_{12}	.04	.05	.00	(bc)	.44	.16
	Average	x_{22}	.69	.03	.71	.06	.80	.05
	Above average	x_{32}	1.00	(bc)	1.00	(bc)	1.00	(bc)
Warranty	90 days	x_{13}	.62	.03	.20	.08	.82	.05
	1 year	x_{23}	.82	.03	.59	.09	.90	.03
	5 years	x_{33}	1.00	(bc)	.80	.08	1.00	(bc)
Service	Ship back to manufacturer	x_{14}	.56	.03	.24	.07	.80	.06
	Service at local dealer	x_{24}	.90	.03	.79	.06	.96	.02
	On-site service	x_{34}	1.00	(bc)	1.00	(bc)	1.00	(bc)
Price	Low	x_{15}	.88	.08	.60	.08	.93	.07
	Medium-low	x_{25}	.77	.07	.44	.08	.88	.07
	Medium-high	x_{35}	.50	.08	.25	.07	.70	.10
	High	x_{45}	.32	.08	.00	(bc)	.62	.12
Choice exponent		β	4.38	.80	2.85	1.40	8.29	2.27
Mixing proportions		w	1.00		.34	.06	.66	.06
Number of respondents			297		101		196	

Notes: bc = boundary condition.

Table 5 shows the parameter estimates for the nested Dogit model. The values do not suggest a noncompensatory consideration process. The estimated captivity parameters are mostly insignificant. Segment 1 consumers have approximately a 30% chance of not considering any alternative. Segment 2 consumers consider almost all alternatives.

Predictions

Consider a scenario in which consumers can choose, at most, one of Brands A, B, and E, each offering above-average performance, a five-year warranty, and on-site service. Let Brands B and E have a medium-low price. We compare the predictions of the two models at different prices for Brand A.

- If Brand A has a high price, its predicted choice probability is .36 by the proposed model and .16 by the nested Dogit model.
- If Brand A has a low price, (1) in Segment 1, Brand C has a predicted choice probability of .35 by the proposed model and .18 by the nested Dogit model, and (2) in Segment 2, Brand E has a predicted choice probability of .26 by the proposed model and .17 by the nested Dogit model.
- The predicted difference in the choice probabilities for Brand A when its own price goes from high to low is .14 by the proposed model and .28 by the nested Dogit model.
- The predicted difference in the choice probabilities for Brand C when the price of Brand A goes from high to low is .09 by the proposed model and .20 by the nested Dogit model.
- The predicted market shares of the brands, which we obtain by dividing the choice probabilities by 1 – *p*(no choice), differ by as much as 15% between the proposed model and the nested Dogit model.
- The predicted no-choice probability for Segment 1 is .01 from the proposed model and .29 from the nested Dogit model. The reason for the difference is that this probability depends on the attribute levels of alternatives in the proposed model but is constant for the nested Dogit model.
- For Segment 1, at all price levels, the nested Dogit model predicts a lower choice probability for Brand A than does the proposed model. For Segment 2, the pattern is reversed. For Brands C and E, in general, the predicted choice probabilities are lower for the Dogit model in both market segments.

These differences in model predictions stem from two sources. First, unlike the nested Dogit model, the proposed choice model allows no-choice probabilities to depend on the attribute levels of the alternatives in a choice set. Second, the proposed model, but not the nested Dogit model, takes into account all possible consideration sets when computing choice probabilities. A priori, we cannot determine which of the two models is more accurate in representing actual choices, but if consumers use noncompensatory rules to evaluate alternatives, these results suggest that linear models are not as robust as we might believe them to be.

CONCLUSION

There is substantial literature on choice-set formation that suggests that people use conjunctive/disjunctive types of rules to screen alternatives before making a choice. We propose subset-conjunctive generalizations of these rules, relax the assumption that consumers use the rules without error and uncertainty about the acceptability of attribute levels, develop a method to infer probabilistic subset-conjunctive rules for unobserved consumer segments, and propose a choice model in which we use consideration probabilities to predict choice probabilities.

We close by noting two possible generalizations. The first is the modeling of multiple screening rules; the satisfaction of any one of these rules classifies an alternative into an acceptable category. The second is an extension of probabilistic subset-conjunctive rules for problems for which there are multiple, possibly ordered response categories.

APPENDIX: REPRESENTING ERROR-FREE, DETERMINISTIC SUBSET-CONJUNCTIVE RULES BY LINEAR MODELS

We show that (1) deterministic subset-conjunctive rules can be represented by linear models when there is no response error and (2) there are linear models with only main effects (and also linear models with only interaction effects) that can represent subset-conjunctive rules. We describe these results for subset-conjunctive rules over m binary attributes. The extension to attributes with arbitrary numbers of levels is straightforward.

Let $x_j = 1(0)$ denote the acceptable (unacceptable) level of the binary attribute j, $1 \le j \le m$. Let

$$\text{(A1)} \qquad y = \begin{cases} 1 & \text{if } u \ge 0, \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\text{(A2)} \qquad u = \beta_0 + \sum_{j=1}^{m} \beta_j x_j + \beta_{1\cdots m} x_1 x_2 \cdots x_m$$

is a linear function with terms representing main effects and the interaction effect among all m attributes. We show that each of the following two parameterizations represents a conjunctive rule.

(Condition A) $\beta_0 = -\alpha\beta$, $\beta_j = \beta$, for all $1 \le j \le m$, and $\beta_{1\ldots m} = 0$, where $m - 1 < \alpha \le m$ and $\beta > 0$.

(Condition B) $\beta_0 = -\beta$, $\beta_j = 0$, for all $1 \le j \le m$, and $\beta_{1\ldots m} = \beta$, where $\beta > 0$.

Note that Condition A specifies a main-effects model and Condition B specifies an interaction-effects model. The claim that both Conditions A and B represent a conjunctive rule implies that the standard interpretations of main and interaction effects do not always make sense when interpreted in the context of quantal response models of the type that Equation A1 represents.

We begin by showing that Condition A represents a conjunctive rule. Substituting the values of the parameters in Condition A into Equation A2 leads to

$$\text{(A3)} \qquad u = -\alpha\beta + \sum_{j=1}^{m} \beta x_j = \beta\left(-\alpha + \sum_{j=1}^{m} \beta x_j\right).$$

As $\beta > 0$, the condition $u \ge 0$ is equivalent to $\Sigma_j x_j \ge \alpha$; thus, Equation A1 becomes

Table 5

DOGIT PARAMETER ESTIMATES FOR THE POOLED (G = 1) AND TWO-SEGMENT SOLUTIONS IN PC STUDY

			Dogit Model					
			Pooled		*Segment 1*		*Segment 2*	
Attribute	*Level*		*Estimates*	*Standard Error*	*Estimates*	*Standard Error*	*Estimates*	*Standard Error*
Brand	B	x_{21}	*–.16*	.1	–.17	.24	–.19	.13
	C	x_{31}	**–.39**	.12	*–.46*	.27	**–.47**	.16
	D	x_{41}	**–.58**	.12	*–.36*	.26	**–.76**	.17
	E	x_{51}	**–1.01**	.12	**–1.30**	.30	**–.97**	.16
Performance	Average	x_{22}	**1.91**	.13	**2.89**	.42	**1.75**	.16
	Above average	x_{32}	**2.71**	.14	**3.52**	.45	**2.68**	.18
Warranty	1 year	x_{23}	**.51**	.08	**1.20**	.23	**.34**	.11
	5 years	x_{33}	**1.00**	.08	**1.86**	.26	**.81**	.11
Service	Service at local dealer	x_{24}	**.83**	.08	**1.41**	.24	**.70**	.11
	On-site service	x_{34}	**1.12**	.08	**2.05**	.27	**.91**	.11
Price	Medium-low	x_{25}	**–.27**	.08	**–.51**	.21	**–.20**	.11
	Medium-high	x_{35}	**–1.02**	.1	**–.92**	.25	**–1.16**	.14
	High	x_{45}	**–1.51**	.11	**–1.77**	.29	**–1.55**	.14
Captivity parameters	Alternative A	γ_1	.00	(bc)	.00	(bc)	.00	(bc)
	Alternative B	γ_2	.00	(bc)	.00	(bc)	.00	(bc)
	Alternative C	γ_3	.02	.01	.02	.02	*.03*	.02
	Alternative D	γ_4	.01	.01	.00	(bc)	*.02*	.02
	Alternative E	γ_5	*.01*	.01	*.02*	.02	.01	.01
No-choice parameter		$1 - \phi$	**.12**	.01	**.29**	.02	.03	.01
Mixing proportions		w	1.00		**.36**	.04	**.64**	.04
Number of respondents			297		108		189	

Notes: We chose the first level of each attribute as the reference in dummy coding; estimates in bold (italics) are significant at $p < .05$ ($p < .1$); bc = boundary condition.

(A4) $$y = \begin{cases} 1 & \text{if } x_1 + \cdots + x_m \geq \alpha, \\ 0 & \text{otherwise.} \end{cases}$$

As $m - 1 < \alpha \leq m$, $y = 1$ if $x_j = 1$ for all $1 \leq j \leq m$. Otherwise, $y = 0$. That is, the conditions in Condition A specify a conjunctive rule.

Next, we show that Condition B also implies a conjunctive rule. Substituting the values of the parameters in Condition B into Equation A2, we obtain

(A5) $$u = -\beta + \beta x_1 x_2 ... x_m = \beta(1 + x_1 x_2 ... x_m).$$

As $\beta > 0$, the condition $u \geq 0$ is equivalent to $x_1, x_2, ..., x_m \geq 1$; thus,

(A6) $$y = \begin{cases} 1 & \text{if } x_1 x_2 \cdots x_m \geq 1; \\ 0 & \text{otherwise.} \end{cases}$$

The condition for $y = 1$ is satisfied only if $x_j = 1$ for all $1 \leq j \leq m$; otherwise, $y = 0$. Thus, Condition B also specifies a conjunctive rule.

We now consider a t-subset-conjunctive rule. It is easy to observe that (1) Condition A specifies a t-subset-conjunctive rule if $t - 1 < \alpha \leq t$ and (2) Condition B provides an interaction-effects representation of a subset-conjunctive rule for $t \geq 2$. The latter uses the following specification of u in which only the tth order interaction-effects terms are included:

(A7) $$u = -\beta + \sum \beta x_{j_1} \cdots x_{j_t},\ \beta > 0,$$

where the summation extends over all possible subsets of t of the m binary variables x_j, $1 \leq j \leq m$.[8] As $\beta > 0$, the condition $u \geq 0$ is equivalent to $\Sigma x_{j1} ... x_{jt} \geq 1$, where the summation extends over all possible subsets of t attributes. Equivalently, $y = 1$ if $x_j = 1$ for at least t of the m attributes.

The preceding discussion implies the following conclusions: First, we can represent deterministic subset-conjunctive rules by imposing constraints on the parameters of a linear model. Second, the linear model representations are not unique. There are multiple main-effects representations and interaction-effects representations of the same rule. The implication is that error-free rules of this variety can be represented by either nonunique main-effects terms or appropriate interaction-effects terms in a latent utility function over the attributes.

[8]For $t = 1$, there is no interaction-effects representation of a subset-conjunctive (i.e., disjunctive) rule. Note that reversing the coding for the response variable y for a disjunctive rule produces an equivalent conjunctive rule.

Using simulated data, we find a similar "degeneracy" in the parameter estimates when we use a binary logit model to fit data from a subset-conjunctive process with small error. When the error is large, the main and interaction effects are not confounded, but it is also not possible to interpret the underlying process as a subset conjunction; we also find similar fit for the binary logit models and for the corresponding subset-conjunctive models.

REFERENCES

Abe, M. (1999), "A Generalized Additive Model for Discrete-Choice Data," *Journal of Business and Economic Statistics*, 17 (3), 271–84.

Alexe, S., E. Blackstone, P.L. Hammer, H. Ishwaran, M.S. Lauer, and C.E. Pothier Snader (2003), "Coronary Risk Prediction by Logical Analysis of Data," *Annals of Operations Research*, 119 (1), 15–42.

Bettman, J.R. (1979), *An Information Processing Theory of Consumer Choice*. Reading, MA: Addison-Wesley.

Boros, E., P. Hammer, and J.N. Hooker (1995), "Boolean Regression," *Annals of Operations Research*, 58, 201–226.

Coombs, C.H. (1951), "Mathematical Models in Psychological Scaling," *Journal of the American Statistical Association*, 46 (256), 480–89.

Crama, Y., P.L. Hammer, and T. Ibaraki (1988), "Cause-Effect Relationships and Partially Defined Boolean Functions," *Annals of Operations Research*, 16, 299–325.

Einhorn, H.J. (1970), "The Use of Nonlinear Compensatory Models in Decision Making," *Psychological Bulletin*, 73 (3), 221–30.

Elrod, T., R.D. Johnson, and J. White (2003), "A New Integrated Model of Noncompensatory and Compensatory Decision Strategies," *Organizational Behavior and Human Decision Processes*, 95 (1), 1–19.

Gaudry, M. and M. Dagenais (1979), "The Dogit Model," *Transportation Research—Part B: Methodological*, 13B (2), 105–112.

Gensch, D.H. (1987), "A Two-Stage Disaggregate Attribute Choice Model," *Marketing Science*, 6 (3), 223–39.

Gilbride, T.J. and G.M. Allenby (2004) "A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules," *Marketing Science*, 23 (3), 391–406.

Roberts, J. and P. Nedungadi, eds. (1995), *International Journal of Research in Marketing*, 12 (1), 1–95.

Swait, J. (2001), "A Non-Compensatory Choice Model Incorporating Attribute Cutoffs," *Transportation Research—Part B: Methodological*, 35 (10), 903–928.

Tversky, A. (1972), "Elimination by Aspects: A Theory of Choice," *Psychological Review*, 79 (4), 281–99.

Verlegh, P.W.J., H.N.J. Schifferstein, and D.R. Wittink (2002), "Range and Number-of-Levels Effects in Derived and Stated Measures of Attribute Importance," *Marketing Letters*, 13 (1), 41–52.

Wedel, M. and W. Kamakura (1999), *Market Segmentation: Conceptual and Methodological Foundations*, 2d ed. Boston: Kluwer Academic.