MARKETING SCIENCE
Vol. 00, No. 0, Xxxxx 2006, pp. 1–16
ISSN 0732-2399 | EISSN 1526-548X | 06 | 0000 | 0001

DOI 10.1287/mksc.1060.0211

Estimating Heterogeneous EBA and Economic Screening Rule Choice Models

Timothy J. Gilbride
Mendoza College of Business, University of Notre Dame, Notre Dame, Indiana 46556, tgilbrid@nd.edu

Greg M. Allenby
Fisher College of Business, Ohio State University, allenby.1@osu.edu

Consumer choice in surveys and in the marketplace reflects a complex process of screening and evaluating choice alternatives. Behavioral and economic models of choice processes are difficult to estimate when using stated and revealed preferences because the underlying process is latent. This paper introduces Bayesian methods for estimating two behavioral models that eliminate alternatives using specific attribute levels. The elimination by aspects theory postulates a sequential elimination of alternatives by attribute levels until a single one, the chosen alternative, remains. In the economic screening rule model, respondents screen out alternatives with certain attribute levels and then choose from the remaining alternatives, using a compensatory function of all the attributes. The economic screening rule model gives an economic justification as to why certain attributes are used to screen alternatives. A commercial conjoint study is used to illustrate the methods and assess their performance. In this data set, the economic screening rule model outperforms the EBA and other standard choice models and provides comparable results to an equivalent conjunctive screening rule model.

Key words: elimination by aspects; consideration sets; attribute screening; noncompensatory decision processes; conjoint analysis; hierarchical Bayes
History: This paper was received September 15, 2005, and was with the authors 1 month for 1 revision; processed by Carl Mela.

1. Introduction

Marketing researchers have drawn on multiple disciplines to describe consumer choice. A persistent theme from behavioral decision research is that respondents are "cognitive misers" who adopt decision strategies and heuristics to cope with complex choice tasks. The elimination by aspects (EBA) model of Tversky (1972) and two-stage screening rule model suggested by Payne (1976) are two theories that focus on how consumers can use the specific attribute levels of choice alternatives to alleviate the cognitive burden of decision making. Although these behavioral decision theories use process- and information-oriented descriptive models of behavior, they can be understood in terms of a structural economic model given the appropriate objective function and constraints. Estimating these models with heterogeneity, however, has proven difficult when the decision is described in terms of a latent sequence of events or by latent constraints and thresholds. The purpose of this paper is to propose and demonstrate methods for estimating two models of choice motivated by behavioral decision theory—the EBA model and an economic screening rule model, where the screening rule is the result of a trade-off between cognitive effort and expected utility.

The EBA model assumes that consumers eliminate choice alternatives that do not have desired attributes or aspects. The first aspect, or product feature, is selected with a probability proportional to its relative importance, all alternatives without that aspect are eliminated, and the process is repeated until only a single alternative remains. Uncertainty in the choice process comes from the stochastic selection of aspects or attribute levels. The "error terms" in the model therefore arise from attribute levels across choice alternatives, and not from alternative specific error terms. The EBA model has intuitive appeal because the amount of cognitive processing associated with attribute-based choice is considered to be less than with compensatory choice models where consumers consider all the alternatives on all the attributes. The EBA has a closed form but a complicated choice probability, with the complexity increasing as the number of alternatives and aspects used to describe alternatives increases. Published studies of EBA models usually include a small number of attributes and/or alternatives and have not dealt with the issue of respondent heterogeneity.

Screening rule models assume respondents screen out some choice alternatives and then engage in effortful processing of the remaining alternatives. Gilbride and Allenby (2004) propose methods for estimating choice models with conjunctive, disjunctive, and compensatory screening rules, where the first two rules involve thresholds for attribute levels and the latter rule is based on the overall utility of an offering being greater than a threshold value. The conjunctive model, which requires that an alternative be acceptable on all relevant attributes in order to be included in the final choice set, was the best fitting model in their study. This paper extends the conjunctive screening rule model by providing an economic rationale for the attribute-level thresholds. Screening rules can result in a loss of utility when an alternative is eliminated on the basis of a single attribute, while consideration of all attributes might reveal it as the most preferred alternative. The economic screening rule model maximizes the number of screening attributes subject to a constraint on the expected loss in utility.

The EBA and economic screening rule model both involve attribute-based decision processes and require that the selected alternative not have any unacceptable attributes. However, the underlying nature of the decision models is different. In the EBA model, individual attributes are used to eliminate alternatives until only a single one remains. The attribute levels that are used will change from choice occasion to choice occasion as well as the order in which they are used to eliminate alternatives. Both the attribute levels that are used and the ordering can result in different alternatives being selected from identical choice sets; consequently, the EBA model is inherently probabilistic.

In the economic screening rule model, certain attribute levels are used to eliminate alternatives from the choice set, but the remaining alternatives are evaluated using all the attribute levels. The screening attributes do not change from choice set to choice set and the order in which the screening takes place does not matter. Given the consumer's current beliefs about the distribution of choice sets, alternatives, and attribute levels, the economic screening rule is a deterministic choice model. It is only factors known to the consumer but not revealed to the researcher that lead to probabilistic models for economic choice models in general and the economic screening rule model in particular. In short, the error term is alternative specific in these models. The different theoretical source of the error term between the EBA and economic screening rule models leads to different approaches for specifying and estimating the models.

The purpose of this paper is threefold. First, a simulation-based methodology is introduced to estimate the EBA model. This method is consistent with attribute-level error terms and overcomes the difficulties in specifying the choice probability that have limited applications of the EBA model. The simulation is facilitated by using a new method for generating quasirandom numbers, the modified Latin hypercube sampling (MLHS) plan proposed by Hess et al. (2005). The MLHS is explained and compared to standard approaches in the context of estimating EBA choice probabilities.

Second, an economic screening rule model is proposed that provides a cost/benefit rationale to attribute-based screening rules. This model results in fewer parameters than comparable screening rule models and uses standard assumptions from econometric models of discrete choice about the source and distribution of the error term. Heterogeneity is introduced into both models via a hierarchical Bayes structure and Markov chain Monte Carlo (MCMC) methods are detailed in the appendix for estimating the models.

Third, the results of an empirical application using a choice-based conjoint study are presented. The performance of the EBA, economic screening rule, a conjunctive screening rule, and the standard economic choice model (as represented by the multinomial logit model) are compared. In this data set, the economic screening rule model provided the best fit to the data. The empirical application demonstrates that the proposed models can be estimated using data commonly collected in marketing research studies.

The paper proceeds as follows. In the next section, the EBA model is reviewed and the simulation-based estimation methodology is detailed. The following section develops the economic screening rule model and derives the likelihood function using assumptions consistent with standard discrete choice models. Results from simulated data sets are presented showing that the models can identify the correct data generating process. Results from an empirical application are then presented. The paper closes with a discussion and suggestions for extending this research.

2. Estimating the EBA Model

The elimination by aspects (EBA) model proposed by Tversky (1972) is one of a number of choice models offered as counterexamples to the rational choice theory of economics. This stream of literature begun by Simon (1955) focuses on how consumers actually make choices given their limited ability to accumulate, process, and make optimal decisions using all the information available in the marketplace. Bettman et al. (1998) provide an overview of the information processing approach to studying consumer choices. In this section we propose a method to estimate the EBA model with stated preference data.

The EBA model characterizes consumer choice as a latent elimination process. Choice alternatives are described by the attributes or aspects that they possess. A consumer places different levels of importance, or weight, on each of these aspects. The choice process begins with the consumer probabilistically selecting one of the aspects; the probability of selecting a given aspect is proportional to its weight. All alternatives without that aspect are eliminated from the choice set; if only one alternative remains, it is chosen. If more than one alternative remains, a second aspect is chosen, again with probability proportional to its weight relative to the weight of the remaining aspects. The elimination process continues until only one alternative remains.

The following example is adapted from Maddala (1983, p. 65) and is used to illustrate the process and the choice probability. Formal mathematical treatments can be found in Tversky (1972) and Batsell et al. (2003). Suppose there are three alternatives described on seven aspects that are shared by the alternatives:

$$A_1 = \{\alpha_1, \alpha_{12}, \alpha_{13}, \alpha_{123}\}$$
$$A_2 = \{\alpha_2, \alpha_{12}, \alpha_{32}, \alpha_{123}\}$$
$$A_3 = \{\alpha_3, \alpha_{13}, \alpha_{32}, \alpha_{123}\},$$

where α_i denotes the weight, or importance, of the aspect to be estimated. Aspect α_1 is unique to alternative A_1, aspect α_{12} is shared by A_1 and A_2, and so on. The use of aspect α_{123} in the EBA model would not eliminate any alternative; it can be ignored. Alternative A_1 can be selected via a number of possible scenarios: α_1 could be selected as the first aspect and all other alternatives eliminated, α_{13} could be selected, eliminating A_2, and then either α_1 or α_{12} could be selected, eliminating A_3, or α_{12} could be selected, eliminating A_3, and then either α_1 or α_{13} could be selected, eliminating A_2. Let S equal the sum over the six relevant α's. Then

$$P(A_1) = \frac{\alpha_1}{S} + \frac{\alpha_{13}}{S}\frac{\alpha_1+\alpha_{12}}{\alpha_1+\alpha_{12}+\alpha_3+\alpha_{32}} + \frac{\alpha_{12}}{S}\frac{\alpha_1+\alpha_{13}}{\alpha_1+\alpha_{13}+\alpha_2+\alpha_{32}}. \quad (1)$$

As the number of alternatives and aspects increases, the number of parameters and the complexity of specifying (1) for each individual for each choice occasion increases. As a result, there have been few published studies in marketing which have estimated the EBA model.

Equation (1) implies that the stochastic nature of the EBA model arises from the selection of particular aspects across the choice alternatives. Choice probabilities are calculated by effectively integrating over the set of aspect-specific error terms that can result in the choice of a particular alternative from the choice set. This contrasts with standard choice models where choice probabilities are calculated by integrating over the set of *alternative* specific error terms that can result in the choice of a particular alternative.

An issue for marketing researchers is how to specify the aspects. In the above example, α_1, α_2, and α_3 are alternative specific aspects and α_{ij} are aspects shared by alternative i and j. Any collections of choice alternatives can be assigned "general" aspects in this fashion. Batsell et al. (2003) estimate the EBA model using repeated observations from a single subject using a "general" aspects coding for five brands of snack foods. They use observed choice probabilities and show that the differences in these choice probabilities can be linearly related to the parameters in the EBA model. This approach requires the respondent to evaluate all possible choice sets repeatedly to get estimates of the choice probabilities. Interpretation of the shared aspects (e.g., α_{12}) is left to the analyst. Manrai and Sinha (1989) provide an application where perceptual data is collected from respondents and two factor scores for each alternative are used to represent the aspects.

Marketing researchers, however, are frequently interested in the influence of particular attributes such as price, color, size, etc. as opposed to "general" aspects. Fader and McAlister (1990) estimate individual-level EBA models where the attributes are limited to brand and promotional status. In their parameterization, each brand has a specific aspect (brand) and a common attribute (promotional status). Consumers choose either from the set of all acceptable brands or from the subset of brands on promotion. Using either "general" aspects or limiting the number of attributes greatly simplifies evaluation of Equation (1).

In the next section, a Bayesian method of estimating a heterogeneous EBA model is proposed. This method uses a simulation-based method for evaluating the likelihood and does not require the construction and evaluation of Equation (1). It therefore is appropriate for marketing studies such as discrete choice conjoint where there are several choice alternatives described on multiple attributes, each with different levels.

2.1. Model Estimation

The model is estimated by replicating the EBA selection process repeatedly for a given set of parameters and calculating the choice probability. Let k index the set of attributes and j enumerate the levels within the attribute. So, if the attributes are brand, size, and color, $K = 3$; the corresponding attribute levels may be $\{\text{brand A}, \text{brand B}, \text{brand C}\}$, $\{\text{large}, \text{small}\}$, $\{\text{white}, \text{black}\}$, and J_k ranges from 2 to 3. Let α_{kj} be the weight given to attribute k level j. These weights are the parameters of interest and they correspond to distinct attribute levels and are not the "general" aspects defined above. For a given choice occasion, the choice process begins by selecting the first attribute level, denoted as $\{kj\}^1$:

$$\Pr(\{kj\}^1) = \frac{\alpha_{kj}}{\sum_{k=1}^{K}\sum_{j=1}^{J_k}\alpha_{kj}}, \tag{2}$$

where the superscript "1" indicates the first step of the elimination process. All alternatives that do not have attribute level $\{kj\}^1$ are eliminated. All remaining alternatives have the same level of attribute k^1 and, therefore, attribute k^1 no longer influences the selection process. The second attribute level $\{kj\}^2$ is then selected with probability:

$$\Pr(\{kj\}^2) = \frac{\alpha_{kj}}{\sum_{k=1,\,k\neq k^1}^{K}\sum_{j=1}^{J_k}\alpha_{kj}}. \tag{3}$$

The process is repeated until a single alternative remains, designated the selected alternative y_i^*. The maximum number of steps in the selection process is K, the number of attributes. The algorithm is summarized as:

For $t = 1$ to K,

(i) Select attribute level $\{kj\}^t$ with probability $\alpha_{kj}/\sum_{k=1,\,k\notin\{k^{t-1},\,k^{t-2},\ldots\}}^{K}\sum_{j=1}^{J_k}\alpha_{kj}$.

(ii) Eliminate all alternatives that do not have attribute level $\{kj\}^t$.

(iii) If only one alternative i remains, set $y_i^* = 1$; otherwise, increase t and go to (i), where the notation $\{k^{t-1}, k^{t-2}, \ldots\}$ represents the set of previously selected attributes.

The choice probability is given by repeating the algorithm (i) to (iii) many times and calculating the selection frequency. Let $y_{im} = 1$ if alternative i was chosen in choice set m by the respondent. Let $s = 1$ to S where S is the total number of times (i) to (iii) is repeated and $y_{ims}^* = 1$ indicate that alternative i was selected in simulation s. The indicator function $I(y_{im} = y_{ims}^*) = 1$ if the chosen alternative matches the alternative selected by the algorithm on simulation s, and 0 otherwise. Then

$$l_m = \Pr(y_{im} = 1) = \frac{\sum_{s=1}^{S} I(y_{im} = y_{ims}^*)}{S}, \tag{4}$$

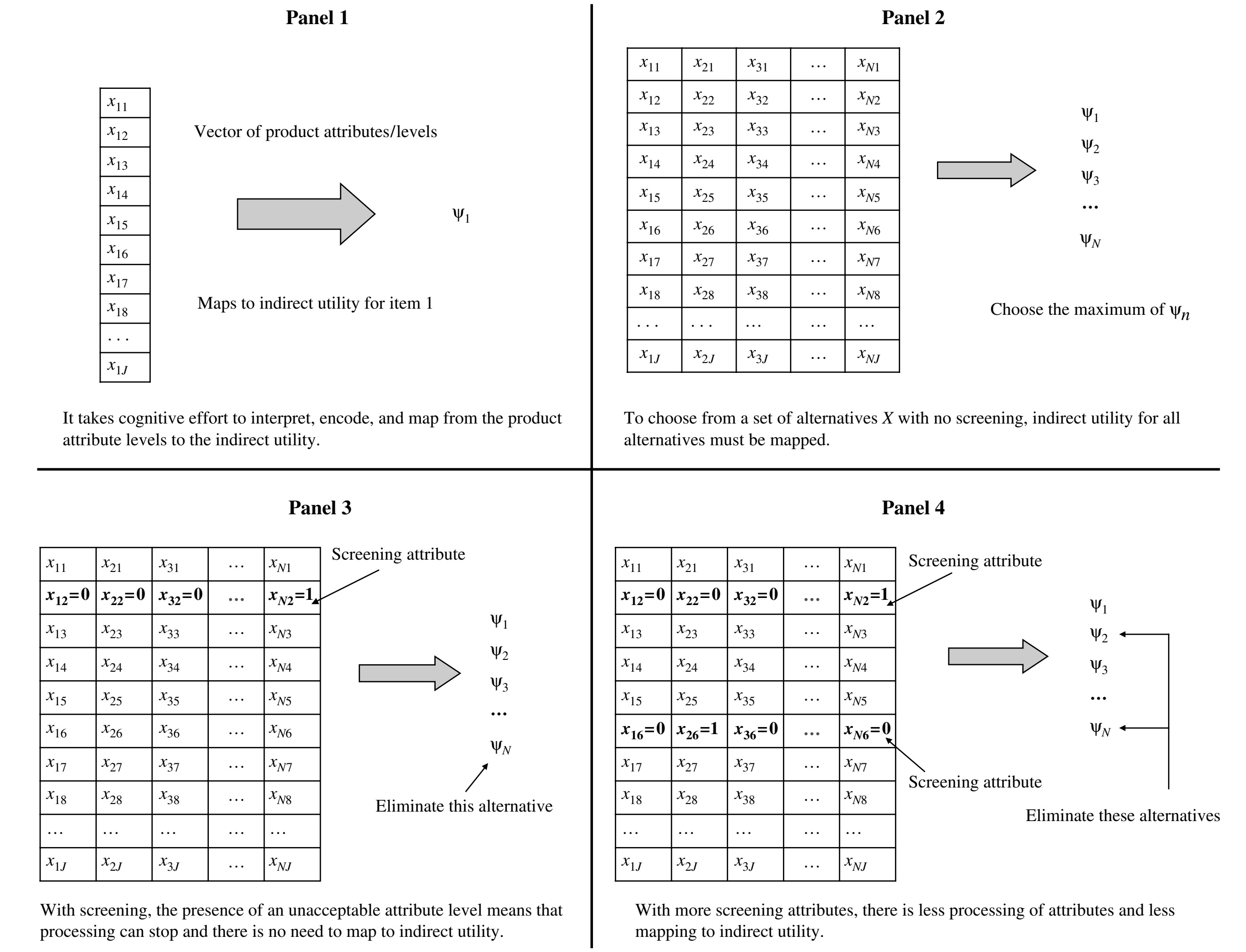
and the conditional likelihood over M choice occasions is given by $l = \prod_{m=1}^{M} l_m$. With the specification of the conditional likelihood, standard MCMC methods can be used to obtain samples from the posterior distributions of the parameters. Appendix A discusses methods of improved estimation of Equation (4) using quasirandom number sequences.

The simulation method outlined in (i) to (iii) and (4) does not require the analyst to specify all the possible paths that could lead to $y_{im} = 1$. Repeating the simulation many times by randomly selecting $\{kj\}^1, \{kj\}^2, \ldots, \{kj\}^K$ effectively integrates over the space of allowable choice paths and, equivalently, attribute-level error terms. Eliminating the need to specify (1) for each respondent in each choice occasion simplifies estimating the EBA model in studies such as conjoint research which involve numerous respondents, alternatives, attributes, and attribute levels. In the above algorithm, if an attribute level that is common to all alternatives is selected as $\{kj\}^t$, then no alternative is eliminated and the algorithm moves on to the next attribute. In some instances, an attribute level may be selected as $\{kj\}^t$ which is not present in any of the remaining alternatives. If "None" is part of the choice set, then it is selected in this instance. If "None" is not part of the choice set, then any attribute levels $\{kj\}$ not present in any of the alternatives are excluded from the selection probability in (i).

A Bayesian approach to modeling heterogeneity is used as compared to previous studies that estimated individual-level fixed effect models. Rossi and Allenby (2003) provide a discussion of the use of Bayesian methods to model heterogeneity. In this application, heterogeneity is introduced by subscripting the importance weights with h signifying a particular household α_{hkj}. The weights must be greater than zero for the selection probability in (i) to make sense so we specify $\alpha_h = \exp(\alpha_h^*)$ and model heterogeneity as $\alpha_h^* \sim \mathrm{N}(\bar{\alpha}^*, V_{\alpha^*})$. Multiplying the weights by a common constant leaves the selection probability in (i) unchanged so the model is unidentified. This is not a problem for Bayesian estimation methods as long as a function of the parameters is identified. In this case, $\Pr(\{kj\}^1)$ is identified and the posterior distribution of $\Pr(\{kj\}^1)$ is reported for each $\{kj\}$. This effectively imposes the identifying restriction $\sum\sum\alpha_{hjk} = 1$ without complicating the MCMC sampler. Postprocessing the draws of the sampler and reporting the identified function of the parameters was suggested by Edwards and Allenby (2003).

The EBA choice model is premised on a rather minimalist behavioral theory: Sequentially eliminate alternatives that do not have the desired aspects or attributes. This basic EBA theory was expanded by Tversky and Sattath (1979) and Gensch and Ghose (1992) to allow aspects to form natural groupings or hierarchies; a consumer then chooses between higher level hierarchies moving down to lower levels of the hierarchy or the "tree" until one alternative remains. A slightly more involved behavioral theory is that a consumer first screens out or eliminates some alternatives and then chooses among the remaining alternatives using a more effortful, compensatory second stage. If the screening rule eliminates alternatives based on particular levels of the attributes, then both the EBA and screening rule model result in a chosen alternative which does not have any unacceptable attributes. Gilbride and Allenby (2004) review

Figure 1 **Illustration of Economic Screening Rule**



two-stage decision processes and provide statistical methods to estimate various screening rule models. In the next section, an economic model is used to determine why certain attributes are used to screen alternatives and how the choice is made from the remaining alternatives.

3. Estimating Economic Screening Rule Models

The economic screening rule model is based on the premise that the presence of an attribute level may indicate that an offering is undesirable and not worth evaluating. Let ψ_i denote the indirect utility for offering i which is identified by the vector of its attribute levels x_i. We assume this vector includes an indicator for log price as an element. The mapping between ψ_i and x_i is represented by $\psi_i(x_i)$. The set of product offerings can be collected into a matrix $X = (x_1, x_2, \ldots, x_n)$ where each column corresponds to a different offering, $1, \ldots, n$. $\psi(X)$ represents the vector where the nth element corresponds to the indirect utility of the nth product offering. We assume that it takes cognitive effort on the part of the decision maker to interpret, encode, and map the matrix X to the vector of indirect utility ψ on any given choice occasion. Panels 1 and 2 in Figure 1 graphically illustrate this mapping from attributes and choice alternatives to indirect utility. This "cost of thinking approach" is discussed in the marketing literature by Hauser and Wernerfelt (1990), Roberts and Lattin (1991), Mehta et al. (2003), and Shugan (1980).

Decision makers can reduce the amount of cognitive effort by limiting the amount of information in X that is evaluated. The presence of any unacceptable element in x_i eliminates the need to further process the other elements in x_i or map it to ψ_i. Panels 3 and 4 in Figure 1 illustrate how the presence of screening attributes reduces the cognitive burden in evaluating specific alternatives x_i and by extension the choice set X. We therefore assume that a consumer minimizes his cognitive effort by maximizing the number of screening attributes. Attribute level "j" is used to screen out, or delete, alternatives if

$$d_j = 1 \quad \text{if } \mathrm{E}_x[\max(\psi(X)) - \max(\psi(X_{-j}))] < \gamma \qquad (5)$$

where $\psi(X)$ is a vector with nth element equal to the indirect utility of choice alternative n, and $\psi(X_{-j})$ is a vector of indirect utilities that excludes all offerings with attribute level j. γ is the amount of indirect utility a decision maker is willing to forego in order to simplify the choice task. γ is inversely related to the amount of cognitive effort the decision maker is willing to spend: as γ approaches zero, no alternatives are screened out and the full matrix X is evaluated. The expectation in Equation (5) is with respect to the respondent's beliefs about the distribution of attribute levels in the market $\pi(x)$. An attribute level is used to screen alternatives if the expectation of the maximum utility of a reduced choice set, excluding offerings with the attribute level, is within γ of the full choice set. The choice model is then:

$$\text{choose alternative } i \text{ if } \psi_i(x_i) > \psi_n(x_n) \quad \text{for all } n \text{ such that } d'x_n = 0 \qquad (6)$$

where d is a vector with elements d_j corresponding to the screening criteria in Equation (5). The screening rule requires that considered alternatives contain no unacceptable attribute levels or, equivalently, that considered alternatives are comprised of only acceptable attribute levels.

The model assumes a consumer utilizes his beliefs about the joint distribution of attribute levels $\pi(x)$ to determine which attributes indicate that an alternative is unacceptable and therefore which alternatives to screen out of any given choice set. Note that an attribute level is "unacceptable" to the extent that it is associated with choice alternatives with relatively low levels of indirect utility.

Anecdotal evidence and the results of Gilbride and Allenby (2004) suggest that consumers may only focus on one or two attributes, such as brand or price, and adopt a screening rule based only on those attributes. The screening rule model implies that this is a reasonable strategy. For instance, if a particular brand is systematically associated with lower levels of other desirable attributes, then the loss of expected utility from excluding that brand from the choice set may be lower than the cost of evaluating all the attributes associated with the brand. Similarly, the psychological benefits of consuming a particular brand may be so low that even at the lowest price and highest levels of all other specified attributes, a consumer may never choose that brand. In either of these situations it makes sense for the consumer to adopt a decision rule that focuses on a particular attribute (e.g., brand) and exclude alternatives with that attribute from the final choice set.

3.1. Model Estimation

Marketing researchers observe the actual choices of a consumer and a potentially incomplete listing of the product attributes. We parameterize the indirect utility function via

$$V_i = \beta' x_i + \varepsilon_i, \qquad (7)$$

and use the error term ε to account for this uncertainty. This mapping from the product attributes to indirect utility and the use of an alternative specific error term is consistent with standard econometric approaches to discrete choice. Equation (5) is now given by

$$d_j = 1 \quad \text{if } \mathrm{E}_x\big[\mathrm{E}_\varepsilon[\max(\beta' X + \varepsilon) - \max(\beta' X_{-j} + \varepsilon)]\big] < \gamma, \qquad (8)$$

where ε is a vector of error terms that we assume to be distributed i.i.d. standard extreme value. The inner expectation is given by the expression (see Anderson et al. 1992)

$$\mathrm{E}_\varepsilon[\max(\beta' X + \varepsilon)] = \ln(\iota' \exp[\beta' X]) + \delta, \qquad (9)$$

where ι is vector of one and δ represents Euler's constant, leading to the screening rule

$$d_j = 1 \quad \text{if } \mathrm{E}_x\big[\ln(\iota' \exp[\beta' X]) - \ln(\iota' \exp[\beta' X_{-j}])\big] < \gamma. \qquad (10)$$

Evaluating Equation (10) requires integrating over the beliefs about the distribution of attribute levels. The empirical study in this paper relies on the pragmatic solution of using the empirical distribution of choice sets faced by the consumer in the study. Thus, the evaluation of Equation (10) proceeds by averaging the expression in brackets on the left side of the inequality over the choice sets in the study. The choice probability is then

$$\Pr(y_i = 1) = \frac{\exp[\beta' x_i]}{\sum_{n=1,\, d' x_n = 0}^{N} \exp[\beta' x_n]}. \qquad (11)$$

Heterogeneity is introduced with a random-effect model for the β and γ parameters $\beta_h \sim \mathrm{N}(\bar{\beta}, V_\beta)$, and because γ_h must be greater than zero, we specify $\gamma_h = \exp(\gamma_h^*)$ and let $\gamma_h^* \sim \mathrm{N}(\bar{\gamma}^*, \sigma_{\gamma^*}^2)$. We use diffuse but proper priors for the hyperparameters. MCMC methods are used to obtain samples from the posterior distributions of all parameters and details are contained in the appendix.

Table 1 reports on a simulation study used to demonstrate the ability to identify the models. Three simulated data sets were generated consisting of 300 heterogeneous respondents, facing 10 choice tasks, each with 6 alternatives. The alternatives were described on four attributes: one with three levels and

Table 1 Log Marginal Density Using Simulated Data

	Model used to estimate parameters		
Data generating process	EBA	MNL	Economic screening rule
EBA	**−4,150.30**	−4,219.40	−4,235.40
MNL	−3,048.40	**−2,964.50**	−2,957.50
Economic screening rule	−2,975.90	−2,963.10	**−2,819.40**

Notes. The log marginal density calculated using the importance-sampling method of Newton and Raftery (1994) is in each cell. The data favors the model with the largest log marginal density.

the remaining attributes at two levels. The choices for the first data set were generated according to the EBA decision model, choices for the second data set were generated according to a multinomial logit (MNL) or the standard economic choice model with extreme value error terms, and choices for the third data set were generated according to the economic screening rule model described in this section. Each data set was used to estimate each of the models, with the true model listed on the vertical dimension of the table and the estimated model along the horizontal dimension. The log marginal density for each model/data set combination is reported in each cell.

The diagonal elements in Table 1 represent the correct model applied to the appropriate data set and indicate that the models can be identified by the data. The EBA model produces the largest log marginal density (LMD) for the EBA data (−4,150.3) compared to the MNL (−4,219.4) and the economic screening rule model (−4,235.4). Similar results apply for the other data sets and models. One anomaly occurs when the screening rule model is used with the MNL data. The LMD for the screening rule model is (−2,957.5) compared to (−2,964.5) for the MNL model with the MNL data. However, note that in the screening rule model, if γ_h is set sufficiently low, all attributes pass the screening rule and no alternatives are screened out. This is the empirical result when the screening rule model was used with the MNL data: The posterior distribution of the parameters implied that virtually no alternatives were screened out of any choice sets (less than 0.7%) and the model recovered the correct hyperparameters for the MNL model. The MNL model is nested in the economic screening rule model (via sufficiently low values of γ_h) and in this example the parameter estimates yielded the correct interpretation.

4. Empirical Application

In this section we describe data and results from a commercial marketing research study using the MNL, EBA, economic screening rule model, and a variant of the conjunctive screening rule model of Gilbride and Allenby (2004).

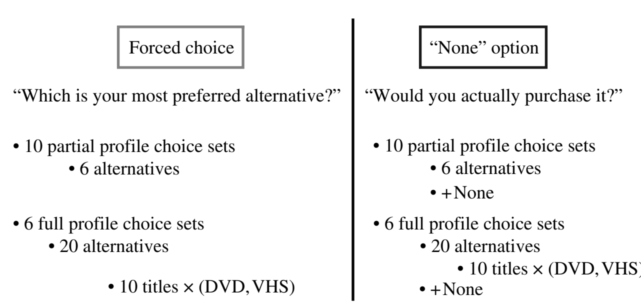
4.1. Data

The proposed models are illustrated using data from a discrete choice conjoint study. Due to the proprietary nature of the data, the actual products are disguised as are the specific product attributes and levels. The study's sponsor is a distributor of documentary films and was interested in consumers' preferences for specific titles and other features of documentaries that are to be sold for the home viewing market. The study involved 10 specific documentaries on different topics and it was thought that only a subset of the titles may appeal to any single consumer. Different prices, types of promotions, and other attributes were included in the study to determine the trade-offs consumers would make. Altogether, seven attributes were included in the design: documentary title, media (DVD versus VHS), packaging, two other disguised attributes, promotion, and price at four different levels. (A complete list of attributes and levels is in Table 4.)

The data collection method included partial profile and full profile choice sets as well as recording choices in the "dual response format." Figure 2 illustrates the different choice sets and the handling of the dual response data. Respondents evaluated a total of 16 choice sets—10 partial profile choice sets which included 6 (of the 10) individual titles described on the remaining 6 attributes, and 6 full profile choice sets with each comprising the 10 documentary titles, each listed in the DVD and VHS format, resulting in 20 alternatives to choose from. The remaining five attributes were also included in the full profile tasks.

Each respondent saw the same 16 choice sets: half evaluated the full profile and then the partial profile, with the order reversed for the other half of the sample. The presentation order for all choice sets was randomized across respondents. For each choice set, consumers were asked to indicate preferred alternatives. In a follow-up question, they were asked if they would actually purchase the selected alternative, if it was available. This is an example of the "dual response format" introduced in commercial marketing research studies to mitigate respondents' proclivity to select the "none" option when it is included as

Figure 2 Dual Response Data and Choice Set Design



one of the choice options. When "none" is selected with too great a frequency, the data becomes noninformative and parameter estimates are unstable.

The purpose of the empirical application is to illustrate the use of the proposed models with data commonly collected in marketing research studies. A variety of different models can be postulated as to how consumers move from indicating the most preferred option to a (hypothetical) purchase decision. Instead of modeling this process, the data is analyzed once using the "forced choice" outcome as the dependent variable and a second time assuming "none" is an alternative in every choice set. In the second analysis, if consumers indicated that they would not purchase the most preferred alternative, then "none" was determined to be the selected alternative. Otherwise, the most preferred alternative was used as the dependent variable. This dual analysis was done to illustrate the models using data collected with and without the "none" option as part of the marketing research study. Parameter estimates from the two analyses are discussed below.

Responses from 296 consumers are available for analysis. In the "forced choice" analysis, consumers made 10 choices from a set of 6 alternatives and 6 choices from a set of 20 alternatives. In the "none" analysis, respondents made 10 choices from a set of 7 alternatives (6 titles + none) and 6 choices from a set of 21 (20 titles + none) alternatives. The "none" option was selected 13.8% of the time. In both treatments of the data, there are a total of $296 \times 16 = 4{,}736$ data points for model estimation.

4.2. Models

Four models are fit separately to data with the "forced choice" outcome and the "none" option. In addition to the EBA and economic screening rule models discussed above, a heterogeneous multinomial logit model and a conjunctive screening rule model are also estimated. Table 2 compares the models in terms of the decision rules, the role and identification of screening attributes, and the error structure.

The conjunctive screening rule model takes the following form:

$$\text{choose alternative } i \text{ if } V_i > V_n \text{ for all } N \text{ such that } I(x_n, \tau) = 1, \quad \text{where} \tag{12a}$$

$$I(x_n, \tau) = 1 \quad \text{if } \Pi_j I(\tau_j > x_{nj}) = 1. \tag{12b}$$

Equation (12a) says to choose the alternative with the maximum indirect utility from those that pass the screening rule. Equation (12b) says that for an alternative to pass the screening rule, it must not have any unacceptable attributes. For dummy coded variables, $x_{nj} \in \{0, 1\}$, then $\tau_j \in \{0.5, 1.5\}$. For $\tau_j = 0.5$, $I(0.5 > 0) = 1$, $I(0.5 > 1) = 0$ and therefore the presence of attribute x_{nj} is unacceptable. Alternatively, for $\tau_j = 1.5$, $I(1.5 > 0) = 1$, $I(1.5 > 1) = 1$ and x_{nj} is an acceptable attribute.

Similar to the model of Gilbride and Allenby (2004), the τ_j's are augmented variables introduced to simplify the analysis. Conditional on the vectors of parameters β and τ, the conditional likelihood is identical to Equation (11) for the economic screening rule model when the same assumptions regarding the error term are made. Heterogeneity is introduced for all parameters and for the augmented τ_{hj} via $\Pr(\tau_{hj} = 0.5) = \theta_j$ modeled across respondents. With binary variables, θ_j is interpreted as the probability that attribute j is unacceptable. As shown in Table 6, this results in 27 values of θ to estimate. Full details of the estimation algorithm are contained in the appendix.

The conjunctive screening rule model is an unrestricted version of the economic screening rule model. Whereas the economic screening rule model implies restrictions on β_h's through the common threshold γ_h (see Equation (10)), the identification of screening attribute levels is unrestricted in the conjunctive

Table 2 Comparison of Models

	Decision rule	Screening attributes	Error structure
Existing models			
Multinomial logit (MNL)	Compensatory model	No screening	Error term assigned to choice alternative
Conjunctive screening rule	Eliminate some alternatives based on attribute levels, choose from remaining using compensatory model	Attribute levels used to screen identified through a parameter on each attribute level	Error term assigned to choice alternative
Proposed models			
Economic screening rule	Eliminate some alternatives based on attribute levels, choose from remaining using compensatory model	Attribute levels used to screen identified through cost/benefit model	Error term assigned to choice alternative
Elimination by aspects (EBA)	Sequential elimination of alternatives based on attribute levels until only one remains	All attribute levels used to eliminate alternatives	"Error term" assigned to attribute level across choice alternatives

model by virtue of a unique τ_{hj} for each attribute level. The restrictions in the economic model arise from the trade-off between expected utility and cognitive effort used for identifying screening attribute levels.

The conjunctive screening rule model has also been modified from the one proposed by Gilbride and Allenby (2004) to match the economic screening rule. Technically, the coding used in Gilbride and Allenby (2004) required an alternative to have *all* the acceptable levels of the attributes in order to pass the conjunctive screening rule. For choice sets with brand names or (in this application) documentary titles, an alternative cannot possess both documentary Title 1 and documentary Title 2. By requiring an alternative to have no unacceptable levels, an alternative with Title 1 or Title 2 could pass the screening rule whereas Title 3 could be used to eliminate alternatives. This formulation of the model not only matches the economic screening rule but addresses several of the limitations noted by Jedidi and Kohli (2005) of the conjunctive screening rule.

Price is treated as a continuous variable in the indirect utility function (via $\ln(\text{Price}_i)$) but as an ordinal variable in the EBA and in the screening rules. In the EBA model, it is assumed that consumers select price based on a "less than or equal to" criteria, e.g., when a consumer selects alternatives that are "less than or equal to Price 2," any alternative with a price "greater than Price 2" is eliminated. This treatment was suggested by Rotondo (1986). In the screening rule model, it is assumed that consumers screen prices based on a "greater than or equal to" criteria, e.g., alternatives "greater than or equal to Price 2" are unacceptable, etc. For the conjunctive screening rule model, this implies the threshold for price is distributed multinomial across respondents. For more information on coding variables for the conjunctive screening rule model see the explanation and examples in Gilbride and Allenby (2004).

The models were estimated using MCMC methods employing hybrid chains that used standard techniques. The EBA and MNL models were estimated using 20,000 iterations with a sample of every tenth iteration from the last 10,000 used to describe the posterior moments. The conjunctive screening rule model was estimated from a run of 50,000 iterations with a sample of every 10th from the final 10,000 used for inference. The economic screening rule model was run for 100,000 iterations again with a sample of every 10th from the final 10,000 used for inference. The economic screening rule model was run for a relatively large number of iterations because simulation studies showed that the parameter $\bar{\gamma}^*$ could be slow to converge; this was not a problem in the empirical applications, however. Convergence was assessed by starting the chains from multiple starting points and inspecting the time series plots of the parameters and identified functions of the parameters. All models in each data set were estimated twice, once with all 16 choice sets per respondent and once with 15 choice sets per respondent used to estimate the model and one used for holdout validation. For each respondent, 1 of the 10 partial profile choice sets was randomly selected for the holdout sample.

4.3. Results

Table 3 contains the fit statistics and predictive results for the four models and the two different methods of handling the dual response format. As in the simulated data sets, the log marginal density (LMD) is computed using the importance sampling method of Newton and Raftery (1994). The hit probability for the holdout sample is the posterior mean of the predicted probability for the selected alternative averaged across iterations of the MCMC chain and respondents. Given the large number of individual-level parameters and the disparity between different models (e.g., 18 for the MNL, 45 in the conjunctive screening rule, 19 in the economic screening rule, and 26 in the EBA model), model fit and parameter estimates are based on results from the full data set. Predictive results are presented as confirmatory evidence and to illustrate similarities in the models.

The economic screening rule model has the highest LMD for both the "forced choice" and "none" treatment of the data and is the favored model for

Table 3 Model Results

	Forced choice			"None" option		
	Full data set	With holdout sample		Full data set	With holdout sample	
	LMD[1]	LMD	Hit prob.[2]	LMD[1]	LMD	Hit prob.[2]
MNL	−4,331.7	−4,085.8	0.487	−4,622.7	−4,442.5	0.469
Conjunctive screening rule	−4,207.6	−3,978.8	0.496	−4,526.4	−4,361.2	0.482
Economic screening rule	**−4,176.3**	**−3,946.8**	**0.497**	**−4,470.0**	**−4,357.7**	**0.483**
EBA	−4,912.0	−4,663.1	0.456	−6,919.2	−6,493.8	0.302

[1]The log marginal density calculated using the importance-sampling method of Newton and Raftery (1994).
[2]Posterior mean of predicted probability for selected alternative.

this data. The fit of the conjunctive screening rule model is similar to that of the economic screening rule model and predictive results are identical. This is not surprising given that the conjunctive model is an "unrestricted" version of the economic screening rule model.

The modest improvement in hit probability for the screening rule models over the standard compensatory model is consistent with other published studies (Gilbride and Allenby 2004, Jedidi and Kohli 2005). As the parameter results show, screening of alternatives was largely limited to documentary titles. When this occurs in models with heterogeneity and a compensatory component, the "brand specific intercepts" can adjust for each respondent to effectively mimic a screening rule process. As demonstrated by the simulation results in Table 1, however, the LMD and parameter estimates from the economic screening rule model are able to differentiate between a compensatory and screening rule model.

The EBA model did not fit this data set as well as any of the other models although it did relatively better with the "forced choice" than with the "none" option treatment of the data. Predictive performance was also below that of the other models. Nonetheless, it is informative to look at the parameter estimates.

The EBA model provides an easy way to compare the relative importance of attributes and levels. Table 4 provides the posterior means (averaged across respondents and iterations of the MCMC chain) of the $\Pr(\{kj\}^1 = 1)$, that is, the probability that attribute k, level j is selected as the first attribute level used to eliminate alternatives. Focusing on the "none" option values, the posterior mean for Title 1 being selected as the first attribute level is 0.036 compared to 0.057 for Title 2; this suggests Title 2 is favored on average to Title 1. Similarly, with a posterior mean of 0.146, being below Price 2 is the attribute level most likely to be chosen first to eliminate alternatives from the choice set. By summing $\sum_j(\Pr(\{kj\}^1 = 1))$ we can measure the relative importance of the attributes. The sum of the posterior means across the documentary titles is 0.644 compared to 0.156 for the different price points and 0.133 for the two types of media. The remaining attributes and attribute levels have relatively low posterior means. As we move from the "forced choice" to the "none" treatment of the data, we find that the price points are relatively more important in the "forced choice" data and, while the importance of "media" is about the same, the relative importance of VHS versus DVD is reversed.

Parameter estimates and functions of parameters for the economic screening rule model are presented in Table 5. With the exception of ln(Price), dummy level coding was used to represent the values of the attribute levels in the indirect utility function. In the "forced choice" treatment of the data, identification was accomplished by setting $\beta_{\text{Title 1}} = 0$. This resulted in 18 values of β_{hj} and one γ_h to be estimated for each individual. In data with a "none" option, identification is achieved by setting the explanatory variables for the "none" option all equal to zero. This allows estimation of $\beta_{\text{Title 1}}$ and results in 19 values of β_{hj} and 1 γ_h to be estimated for each individual. In the "none" treatment of the data, all price points are greater than or equal to the lowest price so it is not used as a screening attribute level. The posterior means and posterior standard deviations of selected parameters from the distribution of heterogeneity are presented in the table. Note that because different identification schemes are used, the values for $\bar{\gamma}^*$, $\sigma^2_{\gamma^*}$, and $\bar{\beta}$ are not directly comparable across the data sets.

The columns labeled "% Unacceptable" are calculated by using the individual values of β_h and γ_h

Table 4 **Elimination by Aspects Model Function of Parameters:** $\Pr(\{kj\}^1)$

	Forced choice		"None" option	
	Posterior mean	Posterior std. dev.	Posterior mean	Posterior std. dev.
Documentary				
Title 1	0.031	(0.003)	0.036	(0.002)
Title 2	0.046	(0.003)	0.057	(0.004)
Title 3	0.027	(0.002)	0.034	(0.003)
Title 4	0.078	(0.004)	0.083	(0.004)
Title 5	0.104	(0.004)	0.099	(0.004)
Title 6	0.048	(0.003)	0.051	(0.003)
Title 7	0.065	(0.004)	0.079	(0.004)
Title 8	0.035	(0.003)	0.036	(0.002)
Title 9	0.043	(0.003)	0.057	(0.003)
Title 10	0.113	(0.005)	0.111	(0.005)
Media				
DVD	0.072	(0.005)	0.055	(0.006)
VHS	0.040	(0.009)	0.078	(0.010)
Packaging				
Packaging 1	0.010	(0.002)	0.013	(0.003)
Packaging 2	0.001	(0.000)	0.001	(0.001)
Attribute 4				
Option 1	0.002	(0.001)	0.006	(0.002)
Option 2	0.017	(0.003)	0.015	(0.003)
Attribute 5				
Option 1	0.000	(0.000)	0.000	(0.000)
Option 2	0.013	(0.003)	0.017	(0.002)
Promotion				
None	0.001	(0.000)	0.001	(0.000)
Promo 1	0.007	(0.001)	0.007	(0.001)
Promo 2	0.001	(0.000)	0.002	(0.001)
Promo 3	0.001	(0.001)	0.003	(0.001)
Promo 4	0.002	(0.001)	0.003	(0.001)
Price				
≤Price 1	0.080	(0.007)	0.005	(0.004)
≤Price 2	0.093	(0.007)	0.146	(0.011)
≤Price 3	0.070	(0.012)	0.005	(0.002)
≤Price 4				

Table 5 Economic Screening Rule Model Selected Parameter Estimates and Functions of Parameters

	Forced choice		"None" option	
	$\bar{\beta}$	(%) Unacceptable	$\bar{\beta}$	(%) Unacceptable
Documentary				
Title 1	0 (0.00)	43.6 (0.04)	3.927 (0.38)	51.0 (0.03)
Title 2	1.236 (0.29)	25.0 (0.03)	4.985 (0.36)	32.8 (0.03)
Title 3	0.615 (0.30)	33.4 (0.04)	4.527 (0.35)	39.0 (0.04)
Title 4	0.908 (0.29)	33.6 (0.03)	4.700 (0.35)	38.0 (0.03)
Title 5	1.220 (0.36)	35.1 (0.03)	5.001 (0.40)	39.6 (0.03)
Title 6	−0.401 (0.33)	44.9 (0.03)	3.511 (0.40)	51.5 (0.03)
Title 7	2.100 (0.25)	15.7 (0.03)	5.716 (0.37)	23.4 (0.02)
Title 8	−0.429 (0.38)	55.0 (0.04)	3.259 (0.45)	61.5 (0.03)
Title 9	0.696 (0.30)	31.4 (0.03)	4.572 (0.37)	38.4 (0.04)
Title 10	2.330 (0.28)	17.5 (0.03)	5.905 (0.39)	23.9 (0.02)
Media				
DVD	−0.139 (0.14)	1.8 (0.01)	−0.143 (0.14)	5.0 (0.01)
VHS		1.0 (0.01)		1.9 (0.01)
Packaging				
Packaging 1	0.305 (0.06)	0.0 (0.00)	0.273 (0.07)	0.0 (0.00)
Packaging 2		0.0 (0.00)		0.1 (0.00)
Attribute 4				
Option 1	−0.092 (0.07)	0.0 (0.00)	−0.069 (0.07)	0.1 (0.00)
Option 2		0.0 (0.00)		0.0 (0.00)
Attribute 5				
Option 1	−0.385 (0.07)	0.0 (0.00)	−0.415 (0.07)	0.2 (0.00)
Option 2		0.0 (0.00)		0.0 (0.00)
Promotion				
None		0.3 (0.00)		1.7 (0.01)
Promo 1	0.458 (0.10)	0.0 (0.00)	0.537 (0.10)	0.4 (0.00)
Promo 2	−0.052 (0.10)	0.2 (0.00)	−0.025 (0.10)	1.8 (0.01)
Promo 3	0.136 (0.10)	0.1 (0.00)	0.185 (0.10)	1.2 (0.00)
Promo 4	0.304 (0.10)	0.1 (0.00)	0.304 (0.11)	1.3 (0.01)
ln(Price)	−2.483 (0.15)		−2.742 (0.16)	
≥Price 1				0.0 (0.00)
≥Price 2		0.0 (0.00)		0.9 (0.00)
≥Price 3		0.6 (0.00)		4.2 (0.01)
≥Price 4		11.4 (0.02)		22.2 (0.02)
Cognitive cost parameters				
$\bar{\gamma}$	−4.102 (0.29)		−3.920 (0.23)	
σ_γ^2	0.613 (0.25)		0.956 (0.30)	

Notes. "Unacceptable" is a function of the parameters β and γ. Posterior means (posterior standard deviations) are presented.

together with Equation (10) to determine the acceptable and unacceptable levels of each attribute level. The average is taken across individuals and across draws of the MCMC chain. These numbers are presented in % format to highlight the difference between them and estimates of θ in the conjunctive screening rule model. The "% Unacceptable" columns show that virtually all of the screening is done on the basis of the title of the documentary and the price. Also, as we move from the "forced choice" to the "none" treatment of the data, consumers are more likely to find any attribute/level unacceptable when "none" is part of the choice set. This makes sense; when consumers are forced to choose, they are more likely to settle for alternatives that they otherwise might have screened out.

Table 6 Conjunctive Screening Rule Model Selected Parameter Estimates

	Forced choice		"None" option	
	$\bar{\beta}$	θ	$\bar{\beta}$	θ
Documentary				
Title 1	0 (0.00)	0.259 (0.06)	3.431 (0.43)	0.097 (0.06)
Title 2	0.503 (0.41)	0.065 (0.04)	5.062 (0.34)	0.120 (0.04)
Title 3	0.136 (0.43)	0.156 (0.07)	4.061 (0.34)	0.050 (0.04)
Title 4	0.374 (0.42)	0.105 (0.06)	4.669 (0.36)	0.082 (0.05)
Title 5	0.463 (0.54)	0.051 (0.04)	4.677 (0.44)	0.032 (0.03)
Title 6	−1.400 (0.63)	0.100 (0.06)	3.233 (0.43)	0.098 (0.05)
Title 7	1.438 (0.33)	0.035 (0.03)	5.614 (0.36)	0.036 (0.03)
Title 8	0.292 (0.60)	0.388 (0.08)	4.893 (0.51)	0.433 (0.06)
Title 9	0.360 (0.34)	0.163 (0.05)	4.978 (0.40)	0.196 (0.05)
Title 10	2.037 (0.40)	0.096 (0.03)	5.855 (0.42)	0.026 (0.02)
Media				
DVD	−0.018 (0.14)	0.041 (0.01)	−0.081 (0.14)	0.045 (0.02)
VHS		0.019 (0.01)		0.035 (0.01)
Packaging				
Packaging 1	0.333 (0.07)	0.003 (0.00)	0.317 (0.07)	0.003 (0.00)
Packaging 2		0.003 (0.00)		0.006 (0.00)
Attribute 4				
Option 1	−0.054 (0.08)	0.010 (0.01)	−0.051 (0.08)	0.005 (0.00)
Option 2		0.003 (0.00)		0.004 (0.00)
Attribute 5				
Option 1	−0.398 (0.07)	0.006 (0.00)	−0.424 (0.07)	0.006 (0.01)
Option 2		0.003 (0.00)		0.003 (0.00)
Promotion				
None		0.007 (0.01)		0.007 (0.01)
Promo 1	0.550 (0.10)	0.004 (0.00)	0.703 (0.10)	0.004 (0.00)
Promo 2	−0.042 (0.10)	0.005 (0.01)	0.037 (0.09)	0.006 (0.01)
Promo 3	0.209 (0.11)	0.004 (0.00)	0.324 (0.11)	0.005 (0.00)
Promo 4	0.376 (0.09)	0.006 (0.01)	0.474 (0.09)	0.008 (0.01)
ln(Price)	−2.518 (0.16)		−2.788 (0.16)	
≥Price 2		0.029 (0.01)		0.032 (0.01)
≥Price 3		0.026 (0.01)		0.023 (0.01)
≥Price 4		0.024 (0.01)		0.031 (0.01)
All prices acceptable		0.921 (0.02)		0.914 (0.02)

Notes. Except where noted for price, θ is the probability that the attribute/level is unacceptable. Posterior means (posterior standard deviations) are presented.

Table 6 contains summary statistics of parameters from the conjunctive screening rule model. The models are identified using the same methods as in the economic screening rule model and resulted in 18 and 19 individual-level β_h parameters in the 2 data sets. In each data set, there are 27 augmented values of τ_h resulting in an equivalent number of θ's to describe the distribution of heterogeneity. The θ's are interpreted as the probability that the attribute level is unacceptable across the sample.

The conjunctive screening rule model indicates that documentary title, media, and to a lesser extent price are used to screen alternatives. The implied screening for media is similar for the economic and conjunctive screening rule models, but the economic screening rule model suggests much higher use of documentary titles and price to screen out alternatives. Comparing the "forced choice" to the "none" option treatment of the data, substantive changes in the conjunctive screening rule parameters are seen for Title 1 (0.259 versus 0.097), Title 2 (0.065 versus 0.120), and Title 3 (0.156 versus 0.050) but the remaining parameters are comparable. This contrasts with the economic screening rule results where a general increase in the use of attribute levels to screen out alternatives was observed when "none" was part of the choice set.

The screening rule models provide better in-sample fit and predictive fit than either the standard multinomial logit or EBA model. The economic screening rule model matches the empirical performance of an equivalent conjunctive screening rule model, but provides a cost/benefit rationale as to why certain attributes are used to screen alternatives and results in many fewer parameters to estimate. The methods proposed in this paper allowed comparisons between the standard model, screening rule models, and the EBA model based on similar treatments of heterogeneity, using the same data sets with and without a "none" option, and using specific attribute levels of interest to marketing managers and product designers.

5. Discussion

This paper has presented methods for estimating two theories of choice that use specific attribute levels to alleviate the cognitive burden of the decision maker. The elimination by aspects (EBA) theory postulates a sequential elimination of alternatives by attribute levels until a single one, the chosen alternative, remains. In the economic screening rule model, respondents screen out alternatives with certain attribute levels and then choose from the remaining alternatives using a compensatory function of all the attributes. The direct comparison of these and other choice models has been hampered by the difficulty of estimating them, with parameter heterogeneity, for reasonably sized choice sets. As the number of alternatives and attribute levels increase, specifying the conditional choice probability in closed form as in Equation (1) becomes more challenging. To deal with this, some researchers have described choice alternatives in abstract terms, limited the number of attributes, and/or estimated individual level models with nonrepresentative samples. Our simulation-based method overcomes these limitations and maintains the stochastic component of the model at the attribute level, as implied by the EBA theory, not at the alternative level. We demonstrate our methods using data from a commercial study with choice alternatives described on 26 specific attribute levels, choice tasks comprising many choice alternatives, and the presence of a "none" choice alternative. Consistent with other modeling approaches, heterogeneity is modeled through a hierarchical Bayesian structure.

The economic screening rule model gives an economic justification as to why certain attributes are used to screen alternatives. The model calculates the expected loss in utility of screening out alternatives with a particular attribute level. Choice alternatives with that attribute level are screened out if the loss in expected utility is less than the cognitive cost of the evaluation process. This economic approach to modeling screening rules results in fewer parameters to estimate as compared to an equivalent conjunctive screening rule model. Our empirical study indicates support for this model relative to the standard MNL model, the EBA model, and a model with a conjunctive screening rule.

Several extensions of this study are possible. The implications of using the dual response format are not well-known and it may predispose respondents to make compensatory trade-offs via the framing of the follow-up question "Would you actually buy this alternative?" If true, this would bias against the EBA model and favor the standard economic or multinomial logit model. Further, we did not explore all the practical uses of the parameter estimates such as identifying "consideration sets" or looking for covariates with individual parameters. Such analyses will be of particular use to marketing managers and the methods and models in this paper are appropriate for investigating these issues. Consistent with other research, the screening rule models provided only modest predictive improvement over the compensatory model. As suggested by Cui and Curry (2005), the use of nonmodel-based predictive approaches such as the support vector machine may provide an upper bound on the predictive ability of model-based approaches, and may provide new insights into appropriate model structures.

The empirical application in this study represents a single data point and additional research with many data sets is needed to understand when and under what conditions consumers may adopt different decision strategies. A growing literature suggests that decision makers may adopt different decision strategies depending on the choice context. Differences in the number of alternatives, number of attributes, types of attributes, the need or desire to be more accurate, etc. may all influence the particular strategy used; see Payne et al. (1992), Bettman et al. (1998), and Gourville and Soman (2005) for theories, experimental evidence, and research propositions relating to these issues.

In addition to choice context, different decision makers may simply adopt different choice strategies. In the empirical analysis, it was assumed that all decision makers used the same choice model, e.g., EBA, economic screening rule, conjunctive screening rule, or compensatory. Because the methods introduced in this paper treat heterogeneity similarly and can be estimated on the same data set, the analysis can be extended to allow different decision makers to use different decision models.

Two theoretical extensions of the economic screening rule model are also possible. First, the model conditions on a consumer's current beliefs about the distribution of attributes and indirect utility across choice alternatives. The model could be expanded by including a dynamic component where consumers update their beliefs; see, for instance, Erdem and Keane (1996). Second, in the empirical application, the "none" option was treated as another alternative with null values for the attributes. Additional insight may be provided by explicitly modeling the budget constraint and the choice of the "outside good."

It is hoped that the models and methods proposed in this paper will be used by researchers to investigate choice processes in settings typically encountered in marketing problems. Additional empirical research is needed to understand if consumers adopt a variety of choice models and under what conditions, or if a single parsimonious representation is sufficient to characterize consumer choice.

Appendix A. Estimation of EBA Choice Probabilities

Calculating the choice probability in (4) is facilitated by using the modified Latin hypercube sampling (MLHS) plan as proposed by Hess et al. (2005). Using quasirandom number sequences such as the MLHS can provide comparable results to random number sequences but with many fewer simulations. The order in which attribute levels $\{kj\}^t$ are selected in (i) of the algorithm (see §2) is governed by the draw of uniform $(0, 1)$ random variables across the K attributes and S simulations. The S simulations in Equation (4) approximate:

$$\int_0^1 g(u) f(u)\, du, \tag{A.1}$$

where $g(u)$ is the selection algorithm given by (i) to (iii) and $f(u)$ is a multivariate uniform density of dimension K. One way to estimate Equation (4) is to generate $S \times K$ independent uniform random numbers. This produces an unbiased estimator of $\Pr(y_{im} = 1)$ whose variance and simulation error decreases as S increases.

The advantage to quasirandom number sequences is that they improve the coverage over the area of integration, e.g., the unit interval. Bearing in mind that computer packages generate only pseudorandom numbers, for a finite number of simulations S, it is possible to have the draws "clustered" in the low end of the unit interval or at the high end of the unit interval, etc. As S increases, the uniform distribution is better approximated and the coverage is improved. The MLHS is designed to ensure better coverage for any finite number of simulations by breaking up the "unit hypercube" into evenly spaced intervals. A one-dimensional sequence of length S is created by:

$$\varphi(s) = \frac{s-1}{S}, \quad s = 1, \ldots, S. \tag{A.2}$$

The sequence is then randomly reordered (shuffled). A multidimensional array of size K is constructed by combining the one-dimensional arrays. The final step is to add a pseudorandom number $0 < u < 1/S$ to each element:

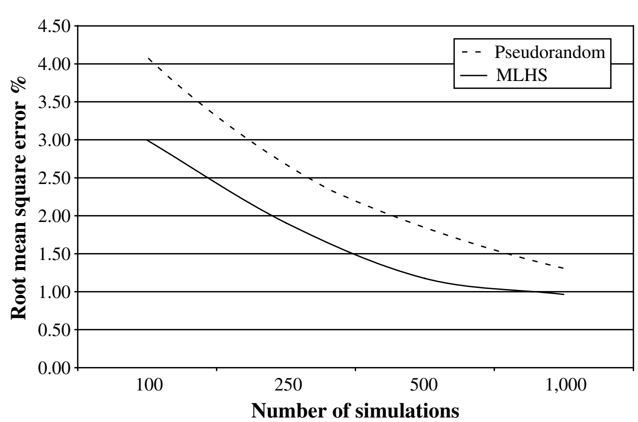
$$\vartheta_k(s) = \varphi_k(s) + u_k \quad s = 1, \ldots, S,\ k = 1, \ldots, K. \tag{A.3}$$

See Hess et al. (2005) for more details and a comparison of the MLHS to other quasirandom plans such as the Halton sequence.

Simulations were conducted to measure the ability of the MLHS plan to estimate EBA choice probabilities. Attribute level weights were assigned and EBA choices were generated for 10 individuals across 20 choice occasions. Each choice set contained 10 alternatives and each alternative was described on 15 binary attributes. Pseudorandom draws were used to generate the actual choices. The actual attribute level weights α_{hkj}, the algorithm (i) to (iii), and Equation (4) were used to determine how well the "true" choice probabilities could be recovered using various sampling plans. The "true" choice probabilities were calculated using pseudorandom numbers and $S = 5{,}000$ for each of the (10 individuals $\times$ 20 choice occasions) $=$ 200 probabilities. This process was repeated 100 times and the average was used as the "true" choice probability.

Simulations using pseudorandom numbers and the MLHS plan with $S = 100$, 250, 500, and 1,000 were then conducted. For example, the choice probability for each of the 200 probabilities was calculated using the MLHS plan with $S = 100$ and the squared difference between the "true" and the predicted probability was recorded. This was replicated 100 times and the average was calculated across replications and choice probabilities to obtain the mean square error. The square root of the mean square errors are presented in Figure 3. As expected, the root mean square error decreases as S increases for both the pseudorandom number and MLHS plans. It is clear, though, that the MLHS has a uniformly lower root mean square error. Based on these

Figure 3 **Recovery of EBA Choice Probabilities MLHS Compared to Pseudorandom Number Generation**



results, $S = 500$ and the MLHS plan were used in the empirical applications.

Appendix B. Estimation Algorithms

This appendix provides detailed algorithms for estimating the models in the paper. Except where noted, the notation follows that introduced in the text.

EBA Model

The model hierarchy is given by

$$y \mid \alpha, X \tag{B.1}$$

$$\alpha^* \mid \bar{\alpha}^*, V_{\alpha^*} \tag{B.2}$$

$$\bar{\alpha}^* \tag{B.3}$$

$$V_{\alpha^*}. \tag{B.4}$$

Equation (B.1) is the likelihood given by (i) to (iii) (see §2) and Equation (4) in the paper where $\alpha = \exp(\alpha^*)$, Equation (B.2) describes the distribution of heterogeneity $\alpha_h^* \sim \mathrm{N}(\bar{\alpha}^*, V_{\alpha^*})$, and Equations (B.3) and (B.4) are conjugate but diffuse priors on the hyperparameters:

$$\bar{\alpha}^* \sim \mathrm{N}(0, 100\Im)$$
$$V_{\alpha^*} \sim \mathrm{IW}(\nu, \Delta).$$

$\Im$ represents the identity matrix and IW is the inverted Wishart distribution with $\nu = V + 8$, $V = 26$ parameters, and $\Delta = \nu I$. The following steps describe an MCMC chain with the posterior distribution of all model parameters as the stationary distribution.

1. *Generate* $\alpha_h^* \mid \bar{\alpha}^*, V_{\alpha^*}, y_h, X$ *for* $h = 1, \ldots, H$. A random-walk Metropolis-Hastings step is used. Let $\alpha_h^{*(p)}$ represent the proposed candidate new draw and $\alpha_h^{*(o)}$ represent the current or old draw. Form $\alpha_h^{*(p)} = \alpha_h^{*(o)} + \eta$ where η is a draw from $\mathrm{N}(0, zV_{\alpha^*})$ and z is a scalar multiple selected to ensure an approximate 50% rejection rate. Form $\alpha_h^{(p)} = \exp(\alpha_h^{*(p)})$.

For $m = 1, \ldots, M$ (choice sets, $M = 16$ in this application):

(a) Generate the MLHS quasirandom numbers (described in Appendix A):

$$\vartheta_k(s) = \varphi_k(s) + u_k \quad s = 1, \ldots, S, \; k = 1, \ldots, K,$$

where $S = 500$ total simulations, and $K = 7$ attributes.

(b) Calculate the simulated choice probability (refer to §2 for first part of algorithm): for $s = 1, \ldots S$ and $t = 1$ to K,

(iv) using $\vartheta_k(s)$, select attribute level $\{kj\}^t$ with probability

$$\frac{\alpha_{hkj}^{(p)}}{\sum_{k=1,\, k \notin \{k^{t-1}, k^{t-2}, \ldots\}}^{K} \sum_{j=1}^{J_k} \alpha_{hkj}^{(p)}}.$$

(v) eliminate all alternatives that do not have attribute level $\{kj\}^t$.

(vi) if only one alternative i remains, set $y_{hims}^* = 1$; otherwise, increase t and go to (i), where the notation $\{k^{t-1}, k^{t-2}, \ldots\}$ represents the set of previously selected attributes:

$$l_{hm}^{(p)} = \Pr(y_{him} = 1) = \frac{\sum_{s=1}^{S} I(y_{him} = y_{hims}^*)}{S}.$$

(c) Accept or reject the proposed vector $\alpha_h^{*(p)}$ with probability:

$$\min: \left(\frac{\prod_{m=1}^{M} l_{hm}^{(p)} \times \exp\left(-\frac{1}{2}(\alpha_h^{(p)*} - \bar{\alpha}^*) V_{\alpha^*}^{-1} (\alpha_h^{(p)*} - \bar{\alpha}^*)\right)}{\prod_{m=1}^{M} l_{hm}^{(o)} \times \exp\left(-\frac{1}{2}(\alpha_h^{(o)*} - \bar{\alpha}^*) V_{\alpha^*}^{-1} (\alpha_h^{(o)*} - \bar{\alpha}^*)\right)}, 1 \right).$$

(d) Postprocess the accepted draws: for $j = 1, \ldots, J_k$ and $k = 1, \ldots, K$,

$$\Pr_h(\{kj\}^1) = \frac{\alpha_{hkj}}{\sum_{k=1}^{K} \sum_{j=1}^{J_k} \alpha_{hkj}}.$$

2. *Generate* $\bar{\alpha}^* \mid \{\alpha_h^*\}, V_{\alpha^*}$.

$$\bar{\alpha}^* \sim \mathrm{N}(\bar{a}, ((V_{\alpha^*}/H)^{-1} + (100I)^{-1})^{-1})$$

$$\bar{a} = ((V_{\alpha^*}/H)^{-1} + (100I)^{-1})^{-1}) \left(V_{\alpha^*}^{-1} \sum_{h=1}^{H} \alpha_h^* + (100I)^{-1}(0) \right).$$

3. *Generate* $V_{\alpha^*} \mid \{\alpha_h^*\}, \bar{\alpha}^*$.

$$V_{\alpha^*} \sim \mathrm{IW}\left(\nu + H, \Delta + \sum_{h=1}^{H} (a_h^* - \bar{\alpha}^*)'(\alpha_h^* - \bar{\alpha}^*) \right).$$

Economic Screening Rule Model

The model hierarchy is given by:

$$y \mid \beta, \gamma, X \tag{B.5}$$

$$\beta \mid \bar{\beta}, V_\beta \tag{B.6}$$

$$\gamma^* \mid \bar{\gamma}^*, \sigma_{\gamma^*}^2 \tag{B.7}$$

$$\bar{\beta} \tag{B.8}$$

$$V_\beta \tag{B.9}$$

$$\bar{\gamma}^* \tag{B.10}$$

$$\sigma_{\gamma^*}^2. \tag{B.11}$$

Equation (B.5) is the likelihood function given in the text by Equation (10), Equations (B.6) and (B.7) are normal distributions for heterogeneity, and Equations (B.8) through (B.11) are conjugate but diffuse priors on the hyperparameters:

$$\bar{\beta} \sim \mathrm{N}(0, 100\Im) \qquad \bar{\gamma}^* \sim \mathrm{N}(0, 100)$$
$$V_\beta \sim \mathrm{IW}(\nu, \Delta) \qquad \sigma_{\gamma^*}^2 \sim \mathrm{IG}(a, b),$$

where the IG represent the inverse gamma distribution, $a = 10$ and $b = 0.1$, and the other notation is as above. The following steps describe an MCMC chain with the posterior distribution of all model parameters as the stationary distribution.

1. *Generate* $\beta_h \mid \gamma_h, \bar{\beta}, V_\beta, y_h, X$ *for* $h = 1, \ldots, H$. A random-walk Metropolis-Hastings step is used. Let $\beta_h^{(p)}$ represent the proposed candidate new draw and $\beta_h^{(o)}$ represent the current or old draw. Form $\beta_h^{(p)} = \beta_h^{(o)} + \eta$ where η is a draw from $\mathrm{N}(0, zV_\beta)$ and z is a scalar multiple selected to ensure an approximate 50% rejection rate. Let j indicate the dummy coded variables used in the indirect utility function: $J = 18$ for "forced choice" and $J = 19$ for "none" option data; let j' indicate the specific attribute levels: $J' = 26$ for "forced choice" and $J' = 27$ for "none" option data (see text and Table 4).

(a) Determine which attributes are unacceptable: for $j' = 1, \ldots, J'$,

$$d_{j'} = 1 \quad \text{if } \frac{1}{M}\sum_{m=1}^{M}\left\{\ln\sum_{n=1}^{N_m}\exp\left(\sum_{j=1}^{J}\beta_{hj}^{(p)}x_{nj}+\beta_{hp}^{(p)}\ln(p_n)\right)\right.$$
$$\left.-\ln\sum_{n^*=1}^{N_m^*}\exp\left(\sum_{j=1}^{J}\beta_{hj}^{(p)}x_{n^*j}+\beta_{hp}^{(p)}\ln(p_{n^*})\right)\right\} < \gamma_h.$$

Otherwise, $d_{j'} = 0$. n^* indicates alternatives without attribute level j'.

(b) Determine which alternatives pass the screening rule and calculate the likelihood: for $m = 1, \ldots, M$ (choice sets), for $n = 1, \ldots, N_m$ (alternatives in choice sets),

$$I_n = 1 \quad \text{if } \sum_{j'=1}^{J'}(d_{j'} \times x_{nj'}) = 0, \quad \text{otherwise } I_n = 0.$$

If $y_{hnm} = 1$ and $I_n = 0$, then reject $\beta_h^{(p)}$, else:

$$l_{hm}^{(p)} = \Pr(y_{him} = 1) = \frac{\exp\left(\sum_{j=1}^{J}\beta_{hj}^{(p)}x_{ij}+\beta_{hp}^{(p)}\ln(p_i)\right)}{\sum_{\substack{n=1\\I_n=1}}^{N_m}\exp\left(\sum_{j=1}^{J}\beta_{hj}^{(p)}x_{nj}+\beta_{hp}^{(p)}\ln(p_n)\right)}.$$

(c) If the proposed vector $\beta_h^{(p)}$ was not rejected in (b), then accept $\beta_h^{(p)}$ with probability:

$$\min: \left(\frac{\prod_{m=1}^{M} l_{hm}^{(p)} \times \exp\left(-\frac{1}{2}(\beta_h^{(p)}-\bar{\beta})V_\beta^{-1}(\beta_h^{(p)}-\bar{\beta})\right)}{\prod_{m=1}^{M} l_{hm}^{(o)} \times \exp\left(-\frac{1}{2}(\beta_h^{(o)}-\bar{\beta})V_\beta^{-1}(\beta_h^{(o)}-\bar{\beta})\right)}, 1\right).$$

2. *Generate* $\gamma_h^* \mid \beta_h, \bar{\gamma}^*, \sigma_{\gamma^*}^2, y_h, X$ *for* $h = 1, \ldots, H$. A random-walk Metropolis-Hastings step is used. Let $\gamma_h^{*(p)}$ represent the proposed candidate new draw and $\gamma_h^{*(o)}$ represent the current or old draw. Form $\gamma_h^{*(p)} = \gamma_h^{*(o)} + \eta$ where η is a draw from $\mathrm{N}(0, z1)$ and z is a scalar multiple selected to ensure an approximate 50% rejection rate; Form $\gamma_h^{(p)} = \exp(\gamma_h^{*(p)})$:

(a) and (b) Follow the steps for drawing β_h using the current β_h and replacing γ_h with $\gamma_h^{(p)}$. $\gamma_h^{(p)}$ is rejected in (b) if it results in a selected alternative being screened out of the choice set.

(c) If the proposed scalar $\gamma_h^{(p)}$ was not rejected in (b), then accept $\gamma_h^{(p)}$ with probability:

$$\min: \left(\frac{\prod_{m=1}^{M} l_{hm}^{(p)} \times \exp\left(-(2\sigma_{\gamma^*}^2)^{-1}(\gamma_h^{*(p)}-\bar{\gamma}^*)\right)}{\prod_{m=1}^{M} l_{hm}^{(o)} \times \exp\left(-(2\sigma_{\gamma^*}^2)^{-1}(\gamma_h^{*(o)}-\bar{\gamma}^*)\right)}, 1\right).$$

Draws of the hyperparameters follow standard conjugate setups and are not detailed here.

Conjunctive Screening Rule
The model hierarchy is given by

$$y \mid \beta, \tau, X \tag{B.12}$$
$$\beta \mid \bar{\beta}, V_\beta \tag{B.13}$$
$$\tau \mid \theta \tag{B.14}$$
$$\bar{\beta} \tag{B.15}$$
$$V_\beta \tag{B.16}$$
$$\theta. \tag{B.17}$$

Equation (B.12) is the likelihood function and Equation (B.13) describes a normal distribution of heterogeneity. $\tau_{hj'}$ are distributed Bernoulli with $\Pr(\tau_{hj} = 0.5) = \theta_{j'}$. For the price variable, $\tau_{h,\text{price}}$ can take on four discrete values and $\tau_{h,\text{price}} \sim \text{multinomial}(H, \theta_{\text{All acceptable}}, \theta_{\geq\text{Price }2}, \theta_{\geq\text{Price }3}, \theta_{\geq\text{Price }4})$. Prior distributions on the hyperparameters are given by

$$\bar{\beta} \sim \mathrm{N}(0, 100\mathfrak{I}) \qquad V_\beta \sim \mathrm{IW}(\nu, \Delta)$$
$$\theta_{j'} \sim \mathrm{Beta}(d, e) \qquad \theta_{\text{price}} \sim \mathrm{Dirichlet}(\delta),$$

where $d = e = 3$ and δ is vector of length 4 with each element set $= 4$. The following steps describe an MCMC chain with the posterior distribution of all model parameters as the stationary distribution.

1. *Generate* $\tau_h \mid \theta, \beta_h, y_h, X$ *for* $h = 1, \ldots, H$. Similar to the algorithm described by Gilbride and Allenby (2004), a "Griddy Gibbs" step is used. Let j indicate the dummy coded variables used in the indirect utility function: $J = 18$ for "forced choice" and $J = 19$ for "none" option data; let j' indicate the specific attribute levels: $J' = 27$ for both data sets (see text and Table 5).

For $j' = 1, \ldots, J' = 23$ (exclude price), $\tau_{hj'}$ can be equal to one of two values: $\tau_{hj'}^{(0.5)} = 0.5$ and $\tau_{hj'}^{(1.5)} = 1.5$. Conditioning on all other τ_h:

(a) Determine which alternatives pass the screening rule and calculate the likelihood for $\tau_{hj'}^{(0.5)}$: for $m = 1, \ldots, M$ and $n = 1, \ldots, N_m$, let $I_n = 1$ if:

$$\Pi_{j'} I\left(\tau_{hj'}^{(0.5)} > x_{nj'}\right) = 1, \quad \text{otherwise } I_n = 0.$$

If $y_{hnm} = 1$ and $I_n = 0$, reject $\tau_{hj'}^{(0.5)}$, else

$$l_{hm}^{(0.5)} = \Pr(y_{him} = 1) = \frac{\exp\left(\sum_{j=1}^{J}\beta_{hj}x_{ij}+\beta_{hp}\ln(p_i)\right)}{\sum_{\substack{n=1\\I_n=1}}^{N_m}\exp\left(\sum_{j=1}^{J}\beta_{hj}x_{nj}+\beta_{hp}\ln(p_n)\right)},$$

$$l_h^{(0.5)} = \prod_{m=1}^{M} l_{hm}^{(0.5)}.$$

(b) Determine which alternatives pass the screening rule and calculate the likelihood for $\tau_{hj'}^{(1.5)}$. This step follows the same procedure as (a), but note that $\tau_{hj'}^{(1.5)}$ will never be rejected as inconsistent with the observed choices. Recall that $\tau_{hj'} = 1.5$ means that both the presence and absence of an attribute is acceptable.

(c) Select $\tau_{hj'}^{(0.5)}$ with probability:

$$\frac{l_h^{(0.5)} \times \theta_{j'}}{l_h^{(0.5)} \times \theta_{j'} + l_h^{(1.5)} \times (1-\theta_{j'})} \quad \text{otherwise, set } \tau_{hj'} = 1.5.$$

For $j' = 24, \ldots, J'$ (the price variable), the multinomial outcomes are selected using a procedure analogous to the one outlined above.

2. *Generate* $\beta_h \mid \tau_h, \bar{\beta}, V_\beta, y_h, X$ *for* $h = 1, \ldots, H$. A random-walk Metropolis-Hastings step is used. Let $\beta_h^{(p)}$ represent the proposed candidate new draw and $\beta_h^{(o)}$ represent the current or old draw. Form $\beta_h^{(p)} = \beta_h^{(o)} + \eta$ where η is a

draw from $N(0, zV_\beta)$ and z is a scalar multiple selected to ensure an approximate 50% rejection rate.

For $m = 1, \ldots, M$ and $n = 1, \ldots, N_m$, let $I_n = 1$ if:

$$\Pi_{j'} I(\tau_{hj'} > x_{nj'}) = 1, \quad \text{otherwise } I_n = 0,$$

$$l_{hm}^{(p)} = \Pr(y_{him} = 1) = \frac{\exp\left(\sum_{j=1}^{J} \beta_{hj}^{(p)} x_{ij} + \beta_{hp}^{(p)} \ln(p_i)\right)}{\sum_{\substack{n=1 \\ I_n=1}}^{N_m} \exp\left(\sum_{j=1}^{J} \beta_{hj}^{(p)} x_{nj} + \beta_{hp}^{(p)} \ln(p_n)\right)}.$$

Accept $\beta_h^{(p)}$ with probability:

$$\min: \left(\frac{\prod_{m=1}^{M} l_{hm}^{(p)} \times \exp\left(-\frac{1}{2}(\beta_h^{(p)} - \bar{\beta}) V_\beta^{-1} (\beta_h^{(p)} - \bar{\beta})\right)}{\prod_{m=1}^{M} l_{hm}^{(o)} \times \exp\left(-\frac{1}{2}(\beta_h^{(o)} - \bar{\beta}) V_\beta^{-1} (\beta_h^{(o)} - \bar{\beta})\right)}, 1 \right).$$

Draws of the hyperparameters follow standard conjugate setups and are not detailed here. See especially Gilbride and Allenby (2004).

References

Anderson, Simon P., Andre de Palma, Jacques-Francois Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. The MIT Press, Cambridge, MA.

Batsell, Richard R., John C. Polking, Roxy D. Cramer, Christopher M. Miller. 2003. Useful mathematical relationships embedded in Tversky's elimination by aspects model. *J. Math. Psych.* **47** 538–544.

Bettman, James R., Mary Frances Luce, John W. Payne. 1998. Constructive consumer choice processes. *J. Consumer Res.* **25** 187–217.

Cui, Dapeng, David Curry. 2005. Prediction in marketing using the support vector machine. *Marketing Sci.* **24**(4) 595–615.

Edwards, Yancy D., Greg M. Allenby. 2003. Multivariate analysis of multiple response data. *J. Marketing Res.* **40**(Aug.) 321–334.

Erdem, Tulin, Michael P. Keane. 1996. Decision making under uncertainty: Capturing dynamic brand choices processes in turbulent consumer good markets. *Marketing Sci.* **15**(4) 1–20.

Fader, Peter S., Leigh McAlister. 1990. An elimination by aspects model of consumer response to promotion calibrated on UPC scanner data. *J. Marketing Res.* **27**(Aug.) 322–332.

Gensch, Dennis H., Sanjoy Ghose. 1992. Elimination by dimensions. *J. Marketing Res.* **29**(Nov.) 417–429.

Gilbride, Timothy J., Greg M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Sci.* **23**(3) 391–406.

Gourville, John T., Dilip Soman. 2005. Overchoice and assortment type: When and why variety backfires. *Marketing Sci.* **24**(3) 382–395.

Hauser, John R., Birger Wernerfelt. 1990. An evaluation cost model of consideration sets. *J. Consumer Res.* **16**(Mar.) 393–408.

Hess, Stephane, Kenneth E. Train, John W. Polak. 2005. On the use of a modified Latin hypercube sampling (MLHS) method in the estimation of a mixed logit model for vehicle choice. *Transportation Res. Part B*.

Jedidi, Kamel, Rajeev Kohli. 2005. Probabilistic subset-conjunctive models for heterogeneous consumers. *J. Marketing Res.* **42**(Nov.) 483–494.

Maddala, G. S. 1983. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, New York.

Manrai, Ajay K., Prabhakant Sinha. 1989. Elimination-by-cutoffs. *Marketing Sci.* **8**(2) 133–152.

Mehta, Nitin, Surendra Rajiv, Kannan Srinivasan. 2003. Price uncertainty and consumer search: A structural model of consideration set formation. *Marketing Sci.* **22**(1) 58–84.

Newton, Michael A., Adrian E. Raftery. 1994. Approximating Bayesian inference with the weighted likelihood bootstrap. *J. Roy. Statist. Soc.* (*B*) **56** 3–48.

Payne, John W. 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organ. Behav. Human Performance* **16** 366–387.

Payne, John W., James R. Bettman, Eric J. Johnson. 1993. *The Adaptive Decision Maker*. Cambridge University Press, New York.

Roberts, John H., James M. Lattin. 1991. Development and testing of a model of consideration set composition. *J. Marketing Res.* **28**(Nov.) 429–440.

Rossi, Peter E., Greg M. Allenby. 2003. Bayesian statistics and marketing. *Marketing Sci.* **22**(3) 304–328.

Rotondo, John. 1986. Price as an aspect of choice in EBA. *Marketing Sci.* **5**(4) 391–402.

Shugan, Steven M. 1980. The cost of thinking. *J. Consumer Res.* **7**(Sept) 99–111.

Simon, Herbert A. 1955. A behavioral model of rational choice. *Quart. J. Econom.* **69**(Feb.) 99–118.

Tversky, Amos. 1972. Elimination by aspects: A theory of choice. *Psych. Rev.* **79**(4) 281–299.

Tversky, Amos, Shmul Sattath. 1979. Preference trees. *Psych. Rev.* **86**(6) 542–573.

Summary of Comments on mksc0211

Page: 1

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:26:47 PM
Type: Note
Provide complete address & confirm e-mail for both authors.

Page: 2

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:27:34 PM
Type: Note
Note which appendix?

Page: 4

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:28:05 PM
Type: Note
"Pr" or "P" (as before) for probability?

Page: 5

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:28:31 PM
Type: Note
OK as set?

Page: 6

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:29:03 PM
Type: Note
Edit OK? define i.i.d.?

Sequence number: 2
Author: Production Editor
Date: 9/20/2006 1:29:18 PM
Type: Note
OK?

Sequence number: 3
Author: Production Editor
Date: 9/20/2006 1:29:32 PM
Type: Note

Which appendix?

Page: 8

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:30:05 PM
Type: Note
Which appendix?

Page: 9

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:31:11 PM
Type: Note
Spell out Prob.?
Head for column 1?

Page: 10

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:33:31 PM
Type: Note
Closing paren, OK?

Page: 12

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:34:14 PM
Type: Note
Ref. list reads Payne et al. (1993)?

Page: 13

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:34:41 PM
Type: Note
Edits OK?

Page: 14

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:37:31 PM
Type: Note
Does this refer to the equation below? If not, please add lead-in sentence if necessary.

Sequence number: 2
Author: Production Editor
Date: 9/20/2006 1:37:58 PM
Type: Note
Add closing parens in fraction.

Page: 15

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:38:36 PM
Type: Note
"z1" as meant?

Sequence number: 2
Author: Production Editor
Date: 9/20/2006 1:39:03 PM
Type: Note
Word missing here?

Sequence number: 3
Author: Production Editor
Date: 9/20/2006 1:39:54 PM
Type: Note
Add closing parens?

Sequence number: 4
Author: Production Editor
Date: 9/20/2006 1:40:31 PM
Type: Note
Greek "Pi" or product?

Sequence number: 5
Author: Production Editor
Date: 9/20/2006 1:42:02 PM
Type: Note
Is this the end of the previous equation? if so, delete period after "1.5." and make "For" lowercase?

Page: 16

Sequence number: 1
Author: Production Editor
Date: 9/20/2006 1:42:21 PM
Type: Note
Greek "Pi" or product?

Sequence number: 2
Author: Production Editor
Date: 9/20/2006 1:42:52 PM
Type: Note
Add closing parens.

Sequence number: 3
Author: Production Editor
Date: 9/20/2006 1:44:52 PM
Type: Note
"2005" in text?
Can you update?
Vol., issue, pages?

Sequence number: 4

Author: Production Editor
Date: 9/20/2006 1:45:41 PM
Type: Note
Cited as "1992" in text?