Category-Based Screening in Choice of Complementary Products

Anocha Aribarg and Natasha Zhang Foutz

Web Appendix

ESTIMATION ALGORITHM

In this section, we will first illustrate how we estimate the proposed model that accounts for heterogeneity in both attribute preferences and use of decision strategies. Our approach to incorporate structural heterogeneity is similar to those in Yang and Allenby (2000), Gilbride and Allenby (2004), and Chib (2001, p.3609). As described in the data section, we conducted a choice-based conjoint study using both single-category and cross-category choice tasks. Let vectors y_{ip}^1 and y_{is}^1, respectively, represent the choices participant *i* made in the nine single-category phone choice tasks, each involving three phones (*p*), and the nine single-category plan choice tasks, each involving three plans (*s*). Similarly, y_i^2 denotes the choices participant *i* made in the first nine of ten cross-category choice tasks. Note that unlike in the single-category choice tasks in which the number of alternatives is fixed at three, the numbers of alternative phone-plan pairs in the cross-category choice tasks vary from task to task depending on the compatible relations between the four phones and four plans in the choice task. We use X_p^1, X_s^1, and X^2, respectively, to represent the attribute-levels of the phones in the single-category phone choice tasks, attribute-levels of the plans in single-category plan choice tasks, and attribute-levels of the compatible phone-plan pairs in the cross-category choice tasks. Heterogeneity in preferences and decision strategies is specified as below:

$$\beta_{ip} \sim MVN(\overline{\beta}_p, \Sigma_{\beta_p}) \qquad \beta_{is} \sim MVN(\overline{\beta}_s, \Sigma_{\beta_s})$$

$$\delta_i \sim Multinomial(\phi_{mb}, \phi_{m1p}, \phi_{m2p}, \phi_{m3p}, \phi_{m1s}, \phi_{m2s}, \phi_{m3s})$$, where

δ_i is an index taking values from 1 to 7, indicating which of the seven possible decision strategies participant i can use, *m* = 1, …, 4 experimental conditions of the conjoint study; *b* = single-stage, *p*=screening by phones, *s*=screening by plans; 1 ~ 3 represents the three cutoffs of interest (1, .66, and .33).

The prior distributions of the hyper-parameters are assumed to be:

$$\overline{\beta}_p \sim MVN(\beta_0^p, \Lambda_{\beta_p}) \qquad \Sigma_{\beta_p} \sim IW(g_p, G_p)$$

$$\overline{\beta}_s \sim MVN(\beta_0^s, \Lambda_{\beta_s}) \qquad \Sigma_{\beta_s} \sim IW(g_s, G_s)$$

$$\phi_m \sim Dirichlet(\alpha_m)$$

A Gibb sampler is constructed as follows.

1. Generate $\beta_{ip} \mid y_{ip}^1, y_i^2, X_p^1, X^2, I(\delta_i, X^2, \beta_{ip}, \beta_{is}), \overline{\beta}_p, \Sigma_{\beta_p}$ using the Metropolis-Hasting. The index δ_i specifies the decision strategy participant *i* uses, and thus how the screening criterion should be specified. For example in our program, if δ_i =1, the screening criterion will be $I\left(\frac{V_{p_k} - V_{\min_p}}{V_{\max_p} - V_{\min_p}} \geq 1\right) = 1$, and only phone-plan pairs that satisfy this criterion will be included in the choice set Ω. β_{ip} are estimated based on $t_1 = 1,...,9$ separate phone choice tasks and $t_2 = 1,...,9$ cross-category choice tasks. We denote the probability of choosing the observed choice *j* in separate phone choice task t_1 as $pr_j^{1t_1}(\beta_{ip})$, and that in cross-category choice task t_2 as $pr_j^{2t_1}(\beta_{ip}, \delta_i)$. For cross-category choice tasks, the probability of choosing observed choice *j* is derived based only on option pairs in the choice set (determined by δ_i), and zero probability is assigned to option *j* if it does not belong to the choice set. This specification assures that new proposed estimates of β_{ip} that give rise to a choice set that does not contain the observed choice *j* in at least one of the joint choice tasks are accepted with probability of zero. That is, β_{ip}^{new} are accepted with probability

$$\min\left(\frac{\prod_{t_1} pr_j^{1t_1}(\beta_{ip}^{new})\prod_{t_2} pr_j^{2t_2}(\beta_{ip}^{new},\delta_i)MVN(\beta_{ip}^{new}-\overline{\beta}_p, D_{\beta_p})}{\prod_{t_1} pr_j^{1t_1}(\beta_{ip}^{old})\prod_{t_2} pr_j^{2t_2}(\beta_{ip}^{old},\delta_i)MVN(\beta_{ip}^{old}-\overline{\beta}_p, D_{\beta_p})},1\right).$$

2. Generate $\beta_{is} \mid y_{is}^1, y_i^2, X_s^1, X^2, I(\delta_i, X^2, \beta_{ip}, \beta_{is}), \overline{\beta}_s, \Sigma_{\beta_s}$ analogously to (1).

3. Generate $\delta_i \mid y_i^2, X^2, \beta_{ip}, \beta_{is}, \phi_m$ where δ_i is an index indicating participant *i*'s decision strategy. To determine δ_i, we need to cycle through our seven possible strategies of interest. Each candidate index $(\delta_i^l), l=1,...7$ is then drawn with probability

$$\frac{\prod_{t_2} pr_j^{2t_2}(\beta_{ip},\delta_i^l)\phi_{lm}}{\sum_k \prod_{t_2} pr_j^{2t_2}(\beta_{ip},\delta_i^k)\phi_{km}}, \; k=1,...,7.$$

Similar to 1, this estimation procedure assures that a strategy that leads to zero probability of choosing observed choice *j* (i.e., choice j does not belong to the choice set) is drawn with zero probability.

4. Generate $\phi_m \mid \{\delta_{im}\}$, for m = 1, …, 4 experimental conditions and M_m is the number of participants in condition *m*:

$$\phi_m \sim Dirichlet\left(\sum_{i=1}^{M_m} I(\delta_i = 1)+\alpha_1, ..., \sum_{i=1}^{M_m} I(\delta_i = 7)+\alpha_7\right).$$

Please refer to Allenby, Arora, and Ginter (1998) for how to generate Dirichlet draws using gamma distribution.

5 ~ 8. Generating $\overline{\beta}_p$ (5), $\overline{\beta}_s$ (6), Σ_{β_p} (7), and Σ_{β_s} (8) is standard. Therefore, we omit the presentation for brevity.

Note that although our goal is to estimate average probabilities of participants engaging in different decision strategies across experimental conditions (posterior means of ϕ_m), we can also average the number of times (all across MCMC draws) that each δ_i^l, l=1,…7 is drawn for participant *i*. These averages can be interpreted as the average probabilities of participant *i* using different decision strategies.

WEB APPENDIX: ADDITONAL INFORMATION ON MODELS 7 AND 8

Model 7 conceptualizes that individuals use enumeration, as opposed to category-based screening, to form choice sets. Nonetheless, a complete enumeration (Andrews and Srinivasan 1995) cannot be directly applied in our study. This is because unlike in the scanner panel data, where there are a relatively small, and often fixed, number of brands, both the *number* (ranging from 7 to 11) and the *identities* of the phone-plan pairs in our study vary across tasks. Thus, it is not only computationally expensive to enumerate all possible choice sets (ranging from 2^7–1 to 2^{11}–1), but more importantly, conceptually meaningless to associate a probability mass to each choice set that has different identities (or compositions) across tasks (Chiang, Chib, and Narasimhan 1999).

We circumvent the *number* problem by constructing the same number of 29 possible choice sets within each task: 14 from enumerating based on 4 unique phones (P1, …, P4, P1/P2, …, P3/P4, P1/P2/P3, …, P2/P3/P4); 14 from enumerating based on 4 unique plans, plus 1 with all available pairs in a task. We circumvent the *identity* problem by further labeling, before enumerating, the four unique phones in each task as, respectively, the best (P1), second best (P2), third best (P3), and worst phone (P4), as determined by estimated part-worths. Similar labeling applies for the four unique plans before they are enumerated. Then a choice set, say {P1}, has a consistent identity, across all tasks, as the choice set that contains only the best phone (and its compatible plans) in a task. Thus, this approach allows meaningful interpretation of the probabilities associated with each of these 29 possible choice sets.

Following a similar approach, we also estimated an alternative version of Model 7 that enumerates all possible choice sets that consist of compatible pairs of the top phones and plans (e.g., a choice set {P2, S3} consists of all compatible pairs of the top two phones and top three plans). The result shows that, similar to Model 7, this model provides better in-sample fit than the proposed model (log marginal density of -1527 and in-sample choice probability of .599) and worse out-of-sample fit (out-of-sample choice probability of .455). Since this model also mimics a strategy of screening by both categories, this result actually provides additional evidence that besides being most appropriate for our research objectives of testing the driving factors of the selection of the screening category, the proposed model of screening-by-one category is capable of describing the observed choices.

As for Model 8, following Model 2 of single-stage compensatory with cross-category interactions, we include the same five interactions in the utility Equation (1).The result shows that such inclusion leads to poorer fit both in-sample and out-of-sample. In fact, none of the estimates related to the interaction terms is significantly different from zero.