Marketing Letters 11:2 (2000): 137–149

A Model for Observation, Structural, and Household Heterogeneity in Panel Data

SHA YANG AND GREG M. ALLENBY
Fisher College of Business, Ohio State University, 2100 Neil Avenue, Columbus, OH 43210

Abstract

Standard methods of understanding customer behavior in marketing allow for differences in sensitivity across consumers, but often assume that the sensitivity of a particular individual is fixed through time. In many situations, this assumption may not be valid. Both the importance of variables, and the manner that they are combined to form an overall measure of value for an offer, can change. In this paper we propose an approach of modeling a customer's purchase history that allows identification of when these aspects of customer behavior are likely to change. This information is useful, for example, in planning when particular customers will be most likely to respond to an offer. Our approach nests common methods of dealing with individual differences, and allows for the introduction of covariates associated with changes in customer behavior. We illustrate our model with data from a national sample of credit card usage and adoption.

Key words: Hierarchical Bayes Models, Direct Marketing, Mixture Distribution

1. Introduction

Consumer preferences and sensitivities change through time. Consumers are price sensitive some days and demand high quality on others. Sometimes they quickly screen out undesired alternatives by using cutoff rules, and other times they have the time and energy to more thoughtfully consider their options. These changes over time occur for both non-durable goods (e.g. buying beer for private consumption versus for a gift) and for goods that last for years (e.g. the evaluation and purchase of an automobile for one's self versus for one's family). The elemental units of analysis in marketing are not just consumers, but consumers within contexts that change through time. It is therefore useful to understand and predict when these changes will occur so that offers made to particular individuals can be tailored to improve response rates and profits.

Current methods of dealing with differences in behavior across consumers do not allow for sufficient flexibility to predict behavioral changes. To be successful, methods need to allow customers to establish their own baseline behavior from which changes can be assessed. This requires models of behavior that allow for both individual differences and temporal dynamics. Models with temporal dynamics typically use covariates such as last brand purchased to explain variation in preferences and sensitivities (e.g. Winer 1986; Bell and Lattin 1996). A limitation of this approach is that consumers with the same covariate values are assumed to have the same sensitivity. Similarly, models that allow for individual differences (Allenby and Rossi 1999) restrict the extent of temporal variation to changes in parameter values, but assume that the same decision rule is being used throughout the consumer's purchase history.

There exists a large body of theoretical and applied research in marketing that establishes the importance of allowing for changes in a consumer's decision rule over time. For example, the literature dealing with the Elaboration Likelihood Model (Cacciopo and Petty 1986) demonstrates that consumers process persuasive arguments by either a central or a peripheral route depending on the degree a decision maker is willing to think about particular actions. Burnkrant and Unnava (1995) demonstrate that these routes can be influenced in advertising by prompting individuals to self-reference, or to consider how they would react in a particular situation. Similarly, customer satisfaction studies often use models of consumer confirmation/disconfirmation (see Cadotte, Woodruff and Jenkins 1987) where expectations and norms are known to vary by factors associated with consumption conditions (see Ostrom and Iacobucci 1995). In these literatures both the variables and decision rules employed by consumers are expected to change over time depending on contextual effects.

Early work by Gensch (1985, 1987) documents the existence of consumers using different decision rules. Models with multiple decision rules are referred to as incorporating structural heterogeneity. This is in contrast to models with preference heterogeneity that assume a common decision process and variables, but allow model coefficients to vary across the decision making units. Recently, Kamakura, Kim and Lee (1996) propose using finite mixture models (see Kamakura and Russell 1989 and Jain, Vilcassim and Chintagunta 1994) to jointly model preference and structural heterogeneity. As stated above, an advantage of modeling structural heterogeneity is that it allows for discrete changes in the likelihood function which can reflect changes in the assumed decision making process. A disadvantage is that these models assume that all observations from a particular unit (e.g. household) are assumed to be generated from the same likelihood.

In this paper we propose a model of observations heterogeneity with the following properties: (1) observations from a particular unit of analysis can come from more than one likelihood; (2) the likelihoods or models can be non-tested, or structurally different; (3) parameters of the data-generating models are continuously distributed across the units in the population; and (4) the probability that a particular observation is generated from a particular model is related to observable covariates. The model allows for observation, structural and household heterogeneity, facilitating the study of covariates associated with changes in behavior. We implement the model in the context of a household's decision to respond to direct mail offers for credit cards, using data from a national panel of credit card usage.

In the next section we describe the model and discuss estimation algorithms that employ Markov chain Monte Carlo methods. We then describe the data and examine the in-sample and predictive fit of various models. The paper concludes with a discussion of potential applications of the methodology.

2. A model for observation heterogeneity

We begin discussion of the model by introducing some general notation. Observations are assumed to be generated according to one of k different models, denoted $f_k(x_{h,k,t}, \beta_{h,k})$

where k indexes the model with covariates x, β denotes model parameters, h indexes the household or decision making unit, and t indexes the individual observations. The k different models can be thought of as representing fundamentally different decision making strategies. Datum from a particular household ($y_{h,t}$) is assumed to come from one of the k different models with probability $\phi_{h,k,t}$. That is,

$$f(y_{h,t}|\{x_{h,k,t}, \beta_{h,k}\}, \gamma_h, z_{h,t}) = \sum_k \phi_{h,k,t}(\gamma_h, z_{h,k,t}) f_k(y_{h,t}|x_{h,k,t}, \beta_{h,k}). \tag{1}$$

Equation (1) is a latent class model (see Lenk and DeSarbo 1997) where each observation of household's data set is assumed generated from model k with probability $\phi_{h,k,t}$. In our implementation of this model we allow for household heterogeneity in β through a continuous random-effect distribution, and allow the latent class probabilities, $\phi_{h,k,t}$, to depend on observable covariates, $z_{h,k,t}$, through a probit specification.

Equation (1) specifies a very general model for studying when consumer sensitivity to an offer may change. When $\beta_{h,k}$ is distributed as a discrete set of point masses and ϕ is indexed only by k, the model is identical to the structural heterogeneity model of Kamakura, Kim, and Lee (1996). When the index of ϕ includes h and k, $\phi_{h,k}$, then heterogeneity is modeled at the observation-level instead of the household-level, but the observation-level probabilities are fixed across observations for any particular unit. By including the covariates $z_{h,k,t}$ in our model specification we allow for the observation-level heterogeneity to vary across observations. These covariates facilitate the description and study of contextual effects, and can be used to predict when changes will likely occur in the nature of consumer demand (f_k).

All of the parameters in equation (1) can not be uniquely estimates, and therefore some restrictions are needed to achieve statistical identification. The model is not identified if both γ_h in $\phi_{h,k,t}$ and $\beta_{h,k}$ in f_k are specified with intercepts. Either specification produces equivalent variation in the product of $\phi_{h,k,t}$ and f_k across a household's observations, and is therefore redundant. The model specification therefore must either allow for intercepts in $\phi_{h,k,t}$ or f_k but not both. In addition, it is important to restrict the amount of variation in the other elements of γ_h so that $\phi_{h,k,t}$, the latent class probability, is not uniquely determined for each observation. If this were to occur then it would not be possible to identify household specific parameters such as $\beta_{h,k}$ because there would be as many parameters as observations. In the analysis reported below we allow for random intercepts in $\phi_{h,k,t}$ and restrict the slope coefficients to be the same across households.

In the next section we describe data from a national survey of credit card usage behavior over a three-year period. Our interest is in building a model that predicts a household's decision to adopt a card by responding to offers received through the mail. We employ a binary probit model of adoption where the probability of responding favorably to an offer at time t is assumed to be equal to:

$$f_k(y_{h,t} = 1|x_{h,t}, r_{h,k,t}, \beta_{h,k}) = \Pr(\text{adopt}_{h,t}|x_{h,t}, r_{h,k,t}, \beta_{h,k}) = \Phi\big((x_{h,t} - r_{h,k,t})' \beta_{h,k}\big) \tag{2}$$

where $\phi(\cdot)$ is the standard normal cumulative density function, $x_{h,t}$ denotes features of the credit card offer at time t and $r_{h,k,t}$ denotes a reference point, or anchor, to which the offer is compared. Differences in the k models correspond to differences in the construction of the reference point $r_{h,k,t}$. Candidates for the reference-point includes features of the cards currently owned by the household, such as the mean interest rate, the minimum annual fee, and so on.

Our illustration deals with whether the reference point is more likely to be card-dependent or feature-dependent for a specific household h. That is, for example, whether the household compares the offer to the best card in their currently held portfolio of credit cards, or whether reference is made to the best features in the portfolio irrespective of the cards. Assuming equal coefficients across latent classes ($\beta_{h,k} = \beta_{h,k'}$), individuals who are more likely to employ card-dependent references are more attractive to firms because new offers are judged relative to a set of attributes which reflect cost constraints in the market.

Heterogeneity within model f_k is then obtained by allowing the model parameters to vary across individuals. We employ a multivariate normal distribution of heterogeneity for each set latent class:

$$\beta_{h,k} \sim \text{Normal}(\overline{\beta_k}, D_k). \tag{3}$$

Finally, we specify the latent class probability for the first class, $\phi_{h,k=1,t}$, as follows:

$$\phi_{h,k=1,t} = \Phi(\gamma_{0,h} + \gamma' z_{h,t}), \tag{4}$$

$$\gamma_{0,h} \sim \text{Normal}(\bar{\gamma}_0, \sigma_\gamma^2). \tag{5}$$

where $z_{h,t}$ are covariates that are believed to be associated with movement across structural models f_k, and $\gamma_{0,h}$ is a household specific intercept that varies across households. As discussed above, the coefficients γ are assumed to be constant across households in this model to achieve statistical identification.

Algorithms for estimating the model are provided in the appendix. The procedure employs a Markov chain Monte Carlo (MCMC) methods (see Gelfand and Smith 1991) to estimate the model parameters. This procedure requires the use of proper prior distributions on the model parameters. In the analysis presented below we assume

$$\overline{\beta_k} \sim \text{Normal } (0, V) \tag{6}$$

$$D_k \sim \text{Inverse Wishart}(G, g) \tag{7}$$

where $V = 400I$, $G = 10I$ and $g = 10$ are parameter specifications on the prior distributions. Because of the large size of the data analyzed below these prior specifications have minimal influence on the posterior parameter estimates.

3. The data

Data were obtained from a national survey of credit card usage. A panel of households was surveyed quarterly for a total of 11 quarters, beginning in April 1994 and running through September 1996. Respondents in the survey were asked to record attributes of their current portfolio of cards, including information about the annual percentage rates (APRs), annual fees, credit limits and the type of card (gold, platinum, etc.). Information also provided credit card usage (e.g. monthly balance) and a variety of socio-economic data describing the respondent and members of their household. In our empirical illustration we assume that the variable $x_{h,t}$ in equation (1) is comprised of the APR and annual fee of the card.

Figure 1 displays time series plots of the average APR and average annual fee for active and new card that had been recently added to the household's portfolio of cards. The average is calculated across all respondents and all cards owned by the respondents. Inactive cards are identified as those that maintain a zero balance and are not used over a three-month period. If a card is used, or if it maintains a positive balance, then the card is considered active. Practitioners commonly use this criterion. New cards are identified by comparing the portfolio of cards reported in one period to the portfolio of cards reported in

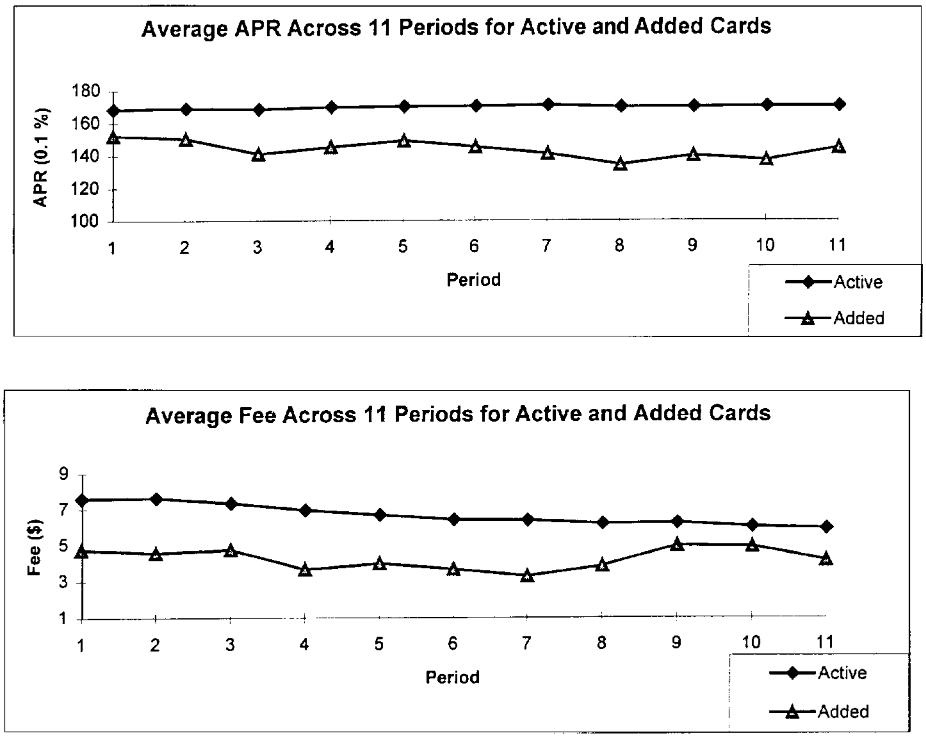

Figure 1. Time series plots of credit card attributes

the previous. The plot shows that, as expected, the added cards have lower APR and fee than the active card currently owned by the respondent.

In the analysis reported below it is necessary to determine the APR and fee of offers in time periods when the respondent elected not to adopt any card. We do this in a manner that is similar to that employed in the analysis of scanner panel data where prices are imputed for the goods not purchased by the household. This imputation is often based on the mean price paid for the items by households during the time period (e.g. during that week). In our data, we know the distribution of APRs and annual fees for new cards adopted in a period, and use the mean values of these variables as explanatory variables when a household does not adopt a card during the same period. In addition, our model assumes that households make a decision to adopt a card in each quarterly period. We feel that this is a reasonable assumption given the high frequency of credit card offers made to households in the sample population (i.e. households in the United States).

Finally, we selected two variables for the vector $z_{h,t}$ in equation (4) above—balance across all cards held in the portfolio in the previous period, and a dummy variable equal to one if the household adopted a new credit card in the previous period. Lagged balance was selected because we believe increases in household debt may be associated with the greater attention paid to the specific features of the credit card offer, while the recent addition of a credit card may lead to a saliency effect where card-based processing dominates. Across panelist, the average range of credit card balances is equal to \$400.

4. Empirical results

The in-sample and predictive fit of a number of models is reported in Table 1, ranging from aggregate models with no heterogeneity to the complete model with context heterogeneity (model 8). The first two models do not include either household or structural heterogeneity, and condition on a portfolio reference point $r_{h,k,t}$. Model 1 assumes the reference point is

Table 1. In-sample and predictive fit

Model	Household Heterogeneity (β)	Structural Heterogeneity (ϕ)	Portfolio Reference	In-Sample Fit[a] (log marginal density)	Predictive Fit[b] (MAD)
1	No	No	Feature	−1049.98	0.329
2	No	No	Card	−1049.64	0.352
3	Yes	No	Feature	−294.24	0.104
4	Yes	No	Card	−326.31	0.123
5	Yes	No	Both Feature and Card	−135.70	0.074
6	Yes	ϕ_k	Feature or Card	−187.36	0.080
7	Yes	$\phi_{h,k}$	Feature or Card	−155.67	0.055
8	Yes	$\phi_{h,k,t}$	Feature or Card	−152.88	0.040

[a]Estimation is based on 1804 observations for 271 households. Reported is the log marginal density of the data calculated using the importance sampling method of Newton and Raftery (1994, p.21).
[b]Predictive fit is based on one observation holdout for 253 households. Only households with 2 more than observations are included in calculating the predictive statistic.

feature-based, while model 2 assumes that the reference point is card-based. The next three models (models 3–5) incorporate household heterogeneity but not structural heterogeneity. Model 3 is the same as model 1 except that household heterogeneity (equation 3) is included in the model specification. Similarly, model 4 is the same as model 2 except for the inclusion of heterogeneity. Finally, model 5 includes both reference points by including four covariates in the model specification—two for the difference between the offer and a feature-based reference point, and two for the difference between the offer and the card-based reference point. Model 5 therefore allows assessment of the gain to allowing for structural and observation heterogeneity beyond that possible with household heterogeneity.

Models six through eight allow for structural heterogeneity in the portfolio reference point. The decision to adopt a card is assumed to come from one of two models that differ in their reference point. Model six assumes that all the observations from a particular household come from one of the two models. Model seven relaxes this assumption by allowing each of the observations to come from one of the two models with probability $\phi_{h,k}$. That is, the likelihood of the model is defined on the individual observations, not the set of observations from a particular household. Finally, model eight introduces covariates, $z_{h,k,t}$, that allows the latent probabilities to vary through time, allowing for the identification and prediction of changes in the decision making model due to contextual effects as reflected by the covariates.

In general, the in-sample fit of the models improves dramatically with the inclusion of household and structural heterogeneity. The log marginal density improves by an order of magnitude across the models, indicating that respondents exhibit large differences in the importance they attach to the variables and how they arrive at their decision. Overall, the fifth model has the best in-sample fit, followed by model eight, which incorporates context heterogeneity. Model five attempts to capture structural heterogeneity by including all variables in the model and allowing different consumers to assign different importance to each through the random-effects distribution for household heterogeneity. In contrast, the structural specification can be more restrictive by not allowing, for example, a consumer to be sensitive to one variable from the first structural model and the second variable from a second structural model. A problem with the fifth model is that the covariates associated with different structural models are sometimes highly correlated, leading to unstable parameter estimates that may not forecast well in non-linear models such as the ones we investigate. Moreover, model five is not able to identify any aspects of structural heterogeneity.

The right-most column of Table 1 reports the predictive fit of the various models as measured in terms of the mean absolute deviation between the predicted probability of choice and the observed choice, coded as a dummy variable. The last observation for households with more than two observations was held out of the analysis, and predictions were made for the 253 households with at least two observations available for estimation. The models with both household and structural heterogeneity tend to fit the data best, and the model that allows structural heterogeneity to vary through time (model 8) is the most predictively accurate. Our proposed model of observation heterogeneity also predicts better than the fifth model, which has better in-sample fit and accommodates heterogeneity by including all variables in the model specification, but only allowing for household heterogeneity.

Table 2 reports parameter estimates for our model of observation heterogeneity (model 8). The upper portion of the table reports estimates of the random-effects error distribution of household heterogeneity (equation 3) for coefficients reflecting the importance of the APR and annual fee when evaluating an offer. The coefficients are similar for the two structural reference points (feature-based and card-based), except for the covariance estimate of the random effects distribution. For feature-based processing we estimate a covariance of -3.411 (correlation $= -0.68$) while for card-based processing we estimate a covariance of 2.501 (correlation $= 0.62$). The positive covariance for card-based processors implies that households tend to be sensitive to either both or neither of the covariates. In contrast, the negative covariance for the feature-based model indicates that households tend to be sensitive to either the APR or annual fee, but not both.

The lower portion of Table 2 reports coefficients for the latent class probabilities (ϕ). Context heterogeneity is modeled by allowing the latent class mixing probabilities to vary both across households and across observations. The variance of the intercept term is 0.238 with standard deviation 0.030 indicating a substantial degree of heterogeneity across

Table 2. Observation heterogeneity model parameter estimates (posterior standard deviations)

	Household Heterogeneity ($\beta_{h,k}$)	
	Feature-Based	Card-Based
Mean ($\bar{\beta}_k$):		
Annual Percentage Rate	-0.310	-0.177
	(0.074)	(0.077)
Annual Fee	-9.523	-6.370
	(1.173)	(0.755)
Variance (D_k):		
Annual Percentage Rate	0.521	0.417
	(0.111)	(0.106)
Annual Fee	48.862	39.275
	(11.976)	(8.063)
Covariance	-3.411	2.501
	(0.957)	(0.763)
	Structural Heterogeneity ($\phi_{h,k,t}$)	
Mean ($\bar{\gamma}$):		
Intercept	-0.0043	
	(0.0642)	
Balance_{t-1}	0.0152	
	(0.0084)	
Adoption_{t-1}	0.0017	
	(0.0714)	
Variance (σ^2_{γ}) :		
Intercept	0.2378	
	(0.0297)	

households. Heterogeneity in the intercept allows the model to reflect household behavior relative to an individual-specific baseline.

More important, we find that the posterior distribution for the lagged balance coefficient to be centered away from zero (97% of its mass is greater than zero), while the coefficient for lagged adoption is centered very close to zero. The lagged balance coefficient allows the latent class probability, ϕ, to vary over time as a household's credit card balance varies. If this coefficient were equal to zero, then model 8 would reduce to model 7 in Table 1. Since the posterior distribution for this coefficient primary has positive support, it indicates that as a household's balance increases, it has a higher likelihood of comparing credit card offers to the best feature of their currently held portfolio, and not to the best card in their portfolio. This result is verified in the predictive tests reported in Table 1 where the inclusion of lagged balance and lagged adoption resulted in improved out-of-sample predictions. Therefore, assuming equal sensitivity to the APR and annual fee, the household with higher balance in the previous period (relative to their own baseline), will be more likely to compare a credit card offering to the best features in their presently held portfolio of cards. Credit card companies should avoid these customers because they are harder to please, and instead focus on customers for who have recently reduced their outstanding balance. It is interesting to note that the influence of these covariates in this data cannot be detected in an aggregate model that does not control for the effects of household and structural heterogeneity. That is, the coefficients for these variables are insignificant when included in models 1 and 2 in Table 1.

5. Discussion

In this paper we develop a flexible model of panel data that allows for variation of a panel member's behavior through time. Our model can incorporate non-nested model structures to represent, for example, different decision rules, and we show how to associate covariates with changes in the probability that particular observations are generated from each latent model. We implement the model within a random-effects specification of preference heterogeneity. Therefore the covariates provide a means of studying when changes in behavior will likely occur, after controlling both for the importance an individual gives to decision attributes and the manner that an overall measure of value is generated.

Our analysis of credit card adoptions provides an illustration of using the proposed procedure. While our example is not exhaustive in terms of the number of variables and possible decision making models used by households, it does provide evidence that observation heterogeneity is present and can be modeled by a finite mixture of sub-models, f_k, in equation 1. In any study the completeness of the empirical investigation is clearly dependent on the quality of data available. We anticipate that as the quality of data continues to improve in the future, improvements will also be realized in the richness and completeness of analyses.

Our analysis was restricted to data collected in diary form, which did not allow us to gain deep insights into the types of decision-making strategies used by the household. As firms begin to reply more on the Internet to collect data, web-based panels will emerge that will substantially lower the cost of conducting surveys. In this environment it will be much easier to gain a deep understanding of various psychological constructs and information processing tendencies of individuals, and relate their variation to personal and marketplace contexts. For example, one application of our methodology lies in the study of the dynamics of reference points used in consumer choice. There has been a strong evidence in the marketing literature that consumers use multiple reference points, obtained from both scanner panel data (Mayhew and Winer 1992; Rajendran and Tellis 1994; Mazumda and Papatla 1999) and in experimental settings (Arora, Bodur and Klein 1999). However, little has been known whether the degree of reliance on memory based reference price and stimulus based reference prices varies across occasions for an individual consumer, and if so what situational factors account for the dynamic variation.

Our model provides a method of identifying when specific consumers will likely be most responsive to offers. We find the predictive fit to improve by allowing for both household heterogeneity and temporal dynamics. At a more fundamental level, the proposed methodology also allows study of why households change their behavior. This can be accomplished by using covariates (z) that have been causally related to behavioral changes in prior studies, and using experimental methods to control for confounding effects. Field studies of this type are commonly conducted in direct marketing to assess the format of various offers. By extending these studies to study changes in behavior over time, firms can better understand and respond to the dynamics of consumer behavior.

Appendix

Estimation Algorithms

1. The model can be estimated using standard MCMC methods (see Allenby, Arora and Ginter 1998, or Rossi, McCulloch and Allenby 1996) if it were known which household observations belong to each of the K latent models. Let $s_{h,k,t}$ be a latent index variable equal to one if observation $y_{h,t}$ is generated from model k. Then

$$pr(s_{h,k,t} = k') \propto \phi_{h,k',t} f_{k'}(y_{h,t} | r_{h,k',t}, x_{h,t}, \beta_{h,k'})$$

   where $\phi_{h,1,t} = \Phi(\gamma_{0,h} + \gamma' z_{h,t})$ and $\phi_{h,2,t} = 1 - \phi_{h,1,t}$. Assignment of $\{s_{h,k=1,\ldots K,t}\}$ is made by evaluating $pr(s_{h,k,t} = k')$ for all K, normalizing to add to one, and generating a draw from this discrete distribution.
2. Given $s_{h,k,t}$, latent continuous variables are generated for each of the probit models: equations (2) and (4). First, for equation (2), denote $p_{h,t}$ as the latent variable. Then: Generate $\{p_{h,t}, h = 1 \ldots, H \text{ and } t = 1, \ldots, T_h\}$

$$f(p_{h,t} | \beta_{h,k}, x_{h,t,k}, s_{h,k,t}, y_{h,t}) \propto \text{Truncated Normal}\big((x_{h,t} - r_{h,k,t})' \beta_{h,k}, 1\big)$$

   where $r_{h,k,t}$ is determined by $s_{h,k,t}$, and if $y_{h,t} = 1$ then $p_{h,t} \geq 0$ and if $y_{h,t} = 0$ then $p_{h,t} < 0$.

3. Generate $\{\beta_{h,k}$, h = 1, ..., H and k = 1, ..., K}

$$f(\beta_{h,k}|\overline{\beta_k}, D_k, s_{h,t,k}) \propto Normal\ (M, \Omega)$$

$$M = (D_k^{-1} + X_{h,k}^{*\prime} X_{h,k}^{*})^{-1}(D_k^{-1}\overline{\beta_k} + X_{h,k}^{*\prime} p_h^{*})$$

$$\Omega = (D_k^{-1} + X_{h,k}^{*\prime} X_{h,k}^{*\prime})^{-1}$$

$X_{h,k}^{*}$ is the data matrix for all observations in which $s_{h,k,t} = 1$ for household h
p_h^{*} is the data vector for all observations in which $s_{h,k,t} = 1$ for household h

4. Generate $\{\overline{\beta_k}$, k = 1, ..., K}

$$f(\overline{\beta_k}|\beta_{h,k}, D_k, V) \propto Normal\ (U_k, W_k)$$

$$U_k = (D_k/H)^{-1} W_k \left(\sum_h \beta_{h,k}/H\right)$$

$$W_k = ((D_k/H)^{-1} + V^{-1})^{-1}$$

$$\mathrm{V} = 400I$$

5. Generate $\{D_k$, k = 1, ..., K}

$$f(D_k|\beta_{h,k}, \overline{\beta_k}, G, g) \propto$$

$$\text{Inverted Wishart}\left(\sum_h (\beta_{h,k} - \overline{\beta_k})(\beta_{h,k} - \overline{\beta_k})' + G, H + g\right)$$

G = 10I, and g = 10, where I indexes identify matrix with dimension 2.

6. Equation 4 latent value are denoted at $q_{h,t}$. The conditional distribution of this latent variable is:

$$f(q_{h,t}|\gamma_{0,h}, \gamma, z_{h,t}, s_{h,k,t}) \propto \text{ Truncated Normal}(\gamma_{0,h} + \gamma' z_{h,t}, 1)$$

If $s_{h,k,t} = 1$ then $q_{h,t} \geq 0$ and if $s_{h,k,t} = 0$ then $q_{h,t} < 0$.

7. Generate $\{\gamma_{0,h}$, h = 1, ..., H}

$$f(\gamma_{0,h}|Z_h, q_h^{*}, \gamma) \propto Normal\ \left(\sum_{t=1}^{T_h} q_{h,t}^{*}/T_h, 1/T_h\right)$$

$$q_{h,t}^{*} = q_{h,t} - \gamma' z_{h,t}$$

8. Generate γ

$$f(\gamma|Z, q^*, \gamma_{0,h}) \propto Normal\ ((Z'Z)^{-1}Z'q^*, (Z'Z)^{-1})$$

$$q^*_{h,t} = q_{h,t} - \gamma_{0,h}$$

9. Generate $\bar{\gamma}_0$

$$f(\bar{\gamma}_0|\gamma_{0,h}, \sigma^2_\gamma) \propto Normal\left(\sum_{h=1}^{H} \gamma_{0,h}/H, \sigma^2_\gamma/H\right)$$

10. Generate σ^2_γ

$$f(\sigma^2_\gamma|\gamma_{0,h}, \bar{\gamma}_0, a, b) \propto \text{Inverted Gamma}\ (A, B)$$

$$A = a + H/2$$

$$B = \frac{2}{\sum_h(\gamma_{0,h} - \bar{\gamma}_0)^2 + 2/b}$$

a = 10 and b = 0.1 are prior specifications.

Acknowledgements

The authors thank Gary Russell and Wagner Kamakura for their helpful comments.

References

Allenby, Greg M., Arora, Neeraj and James L. Ginter. (1998). "On the Heterogeneity of Demand," *Journal of Marketing Research* 35, 384–389.

Allenby, Greg M. and Peter E. Rossi. (1999). "Marketing Models of Household Heterogeneity," *Journal of Econometrics*, forthcoming 89, 57–78.

Allenby, Greg M. and James L. Ginter. (1995). "Using Extremes to Design Products and Segment Markets," *Journal of Marketing Research* 33, 392–403.

Arora, Neeraj, Onur Bodur, Noreen Klein. (1999). "Investigating How Consumers Process Multiple Reference Prices," Working Paper.

Bell, David R. and James M. Lattin. (1996). "Looking for Los Aversion in Scanner Panel Data: The Confounding Effect of Price-Response Heterogeneity," Working Paper #1259, Graduate School of Business, Stanford University.

Burnkrant Robert E. and H. Rao Unnava. (1995). "Effects of Self-Referencing on Persuasion," *Journal of Consumer Research* 22, 17–26.

Cadotte, Ernest R., Robert B. Woodruff and Roger L. Jenkins. (1987). "Expectations and Norms in Models of Consumer Satisfaction," *Journal of Marketing Research* 24, 305–314.

Gelfand, Alan and Adrian Smith. (1990). "Sampling-based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Society* 85, 398–409.

Gensch, H. Dennis. (1985). "Empirically Testing a Disaggregate Choice Model for Segments," *Journal of Marketing Research* 462–467.

Gensch, H. Dennis. (1987). "Empirical Evidence Supporting the Use of Multiple Choice Models in Analyzing a Population," *Journal of Marketing Research* 192–207.

Jain Dipak C., Naufel J. Vilcassim and Pradeep K. Chintagunta. (1994). "A Random Coefficients Logit Brand-Choice Model Applied to Panel Data," *Journal of Business and Economic Statistics* 12, 317–328.

Kamakura, Wagner A., Kim, Byung-Do and Jonathan Lee. (1996). "Modeling Preference and Structural Heterogeneity in Consumer Choice," *Marketing Science* 15, 152–172.

Kamakura, Wagner A., and Gary J. Russell. (1989). "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research* 26, 379–390.

Lattin, James M. and Randolph E. Bucklin. (1989). "Reference Effects of Price and Promotion on Brand Choice Behavior," *Journal of Marketing Research* 26, 299–310.

Lenk, Peter J. and Wayne S. DeSarbo. (1997). "Bayesian Inference of Finite Mixture, Random Effects, Generalized Linear Models," working paper, University of Michigan.

Mayhew, Glenn E. and Russell S. Winer. (1992). "An Empirical Analysis of Internal and External Reference Prices Using Scanner Data," *Journal of Consumer Research* June 19(1), 62–70.

Mazumdar, Tridib and Purushottam Papatla. (1999). "An Investigation of Reference Price Segments," Forthcoming, *Journal of Marketing Research*.

Newton, Michael A. and Adrian E. Raftery. (1994). "Approximating Bayesian Inference with the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society (B)* 56, 3–48.

Ostrom Amy and Dawn Iacobucci. (1995). "Consumer Trade-Offs and the Evaluation of Service," *Journal of Marketing* 59, 17–28.

Petty, Richard E. and John T. Cacioppo. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*, New York: Springer-Verlag.

Rajendran, K.N. and Gerald J. Tellis. (1994). "Contextual and Temporal Components of Reference Price" *Journal of Marketing*, 58 (January), 22–34.

Rossi, Peter E., Robert E. McCulloch and Greg M. Allenby. (1996). *The Value of Purchase History Data in Target Marketing*, *Marketing Science* 15, 321–340.

Winer, Russell S. (1986). "A Reference Price Model of Band Choice for Frequently Purchased Consumer Products" *Journal of Consumer Research* 13, 250–56.