

# Beyond the Endogeneity Bias: The Effect of Unmeasured Brand Characteristics on Household-Level Brand Choice Models

Pradeep Chintagunta, Jean-Pierre Dubé, Khim Yong Goh

Graduate School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, Illinois 60637  
{pradeep.chintagunta@gsb.uchicago.edu, jdube@gsb.uchicago.edu, kgoh@gsb.uchicago.edu}

We investigate the role of potential weekly brand-specific characteristics that influence consumer choices, but are unobserved or unmeasurable by the researcher. We use an empirical approach, based on the estimation methods used for standard random coefficients logit models, to account for the presence of such unobserved attributes. Using household scanner panel data, we find evidence that ignoring such time-varying latent (to the researcher) characteristics can lead to two types of problems. First, consistent with previous literature, we find that these unobserved characteristics may lead to biased estimates of the mean price response parameters. This argument is based on a form of price endogeneity. If marketing managers set prices based on consumer willingness to pay, then the observed prices will likely be correlated with the latent (to the researcher) brand characteristics. We resolve this problem by using an instrumental variables procedure. Our findings suggest that simply ignoring these attributes may also lead to larger estimates of the variance in the heterogeneity distribution of preferences and price sensitivities across households. This could overstate the benefits from marketing activities such as household-level targeting. We resolve the problem by using weekly brand intercepts, embedded in a random coefficients brand choice model, to control for weekly brand-specific characteristics, while accounting for household heterogeneity. Overall, our results extend the finding on the endogeneity bias from the mean of the heterogeneity distribution (i.e., the price effect) to include the variance of that distribution.

*Key words:* brand choice; choice models; econometric models; targeting; endogeneity; instrumental variables

*History:* Accepted by Jagmohan S. Raju, marketing; received July 7, 2003. This paper was with the authors 3½ months for 3 revisions.

## 1. Introduction

Researchers estimating logit brand choice models with household scanner panel data have been increasingly concerned with the presence of unmeasured brand characteristics (UBCs) that vary over time and that affect consumer choices. Such characteristics may include shelf space allocated to a brand, its shelf location, and in-store coupons that are specific to the retail environment for a brand in that week and that are common to all households shopping in the store. The UBCs (due to their influence on choices and hence on demand), in turn, would influence retail price decisions, resulting in a correlation between an included variable in the choice model (price) and an unobserved “error” term (the UBC). This correlation generates an endogeneity problem (Berry 1994; Villas-Boas and Winer 1999, which is referred to as V-BW henceforth). Such endogeneity when not properly accounted for results in erroneous estimates for the mean effects of marketing variables (price) on choice behavior. As demonstrated by V-BW, this form of endogeneity is not resolved by using household-level data.

Besides the price endogeneity problem described above, we also expect UBCs, when ignored, to generate overstated variances in the estimated distribution of heterogeneity in household preferences and price sensitivities. Intuitively, this result stems from a variance decomposition argument—if there is a total variance associated with the systematic component of a household’s utility for a brand, then ignoring the UBCs and their associated variance could inflate the estimated variance of the heterogeneity distribution across households. In this study, we provide an empirical approach to account for both forms of bias. The first is the bias associated with price endogeneity, which may affect the mean price response. The second is the bias associated with overstated estimates of the variances in household heterogeneity, which may affect the width of the support of the distribution of price responses. Both sources of bias will have adverse effects on model parameters and on the shape of the distribution of consumer tastes.

The key contributions of this study are as follows: Substantively, unlike previous research, our study documents the impact of ignoring UBCs—not just on

the mean effects of marketing activities (as in V-BW), but also on the distribution of these effects across households as noted above. This is important because marketing policies such as targeting are based on an understanding of the heterogeneity distribution. If the amount of heterogeneity estimated to be present is less after accounting for endogeneity, then the potential profit gains to targeting will be reduced. Methodologically, as compared to previous studies, we make fewer assumptions on the distribution of the UBCs and the nature of their correlation with prices when estimating the parameters of the logit brand choice model (cf. V-BW, Yang et al. 2003). This has two consequences. First, our estimated demand parameters will not suffer from biases that can arise from incorrectly specifying the distribution of the UBCs or their correlation with prices. Second, our methodology alleviates computational difficulties associated with the full-information maximum likelihood-based procedures (discussed later) used previously. Specifically, it can be implemented via a standard random coefficients logit model estimation routine. Furthermore, we are able to allow for heterogeneity in both intercepts as well as slope coefficients in the brand choice model. Not accounting fully for heterogeneity across households runs the risk of finding an endogeneity bias even in its absence, because ignoring either heterogeneity or endogeneity has previously been shown to bias the price parameter in the same direction (i.e., toward zero).

We carry out an extensive Monte Carlo simulation study to compare our approach with alternative full and limited information approaches that have been proposed in the marketing literature.<sup>1</sup> The two key conclusions from the simulation are as follows. First, the proposed approach provides the most robust results across several different mechanisms that can generate the endogeneity problem. Second, consistent with our discussion above, we are able to document the presence of both sources of bias—in the mean as well as the variance of the heterogeneity distribution—when endogeneity is ignored.

To estimate the model parameters, we use a household scanner panel dataset on margarine purchases from the Denver market. Our estimates provide evidence of both forms of bias associated with unmeasured time-varying brand characteristics. First, failure to control for the endogeneity of prices appears to bias the estimated mean price response parameters toward zero. Second, ignoring the role of time-varying characteristics appears to generate substantially larger estimates of the variance in consumer taste parameters. We also find that controlling for unobserved

heterogeneity in brand preferences and the effects of marketing variables leads to similar increased price sensitivity in the estimated demand. Interestingly, accounting for both heterogeneity and endogeneity leads to even more elastic demand, illustrating the importance of incorporating both into the choice model estimation. We then explore the policy implications of these biases. For standard market-structure analyses, using relative price elasticities, we do not find noticeable adverse consequences from ignoring the endogeneity problem. However, if we want to draw inferences based on the absolute magnitudes of the price elasticities, accounting for the UBCs would be very important. We illustrate this point by designing household-specific targeted prices. Our results indicate that the model that accounts for the UBCs predicts lower gains from targeting compared with the standard random coefficients models because, on average, price sensitivity is higher leading to lower prices, and because of the lower dispersion in price sensitivities across households leading to smaller gains from price discrimination.

The remainder of the paper is organized as follows. In §2, we review the literature related to our work. In §3, we discuss the choice model. In §4, we derive the econometric model and explain the estimation procedure. In §5, we describe our data and discuss our empirical results. We then demonstrate how ignoring the endogeneity of prices can affect the posterior distribution of consumer tastes, leading to misleading managerial recommendations both for uniform and targeted pricing policies. Section 5 also discusses our robustness checks. We conclude in §6.

## 2. Literature Review

The primary distinction made by the previous literature in this area involves the treatment of the supply side. Modeling the equilibrium conditions of an economic model of pricing (i.e., the supply side) allows us to control for the potential endogeneity due to UBCs. This approach also provides information for identifying the structural parameters. However, capturing the institutional properties of pricing in a simple economic model may be difficult, and imposing the wrong supply-side model will contaminate the estimates for the demand parameters. For this reason, one may prefer a more agnostic “reduced-form” approach. Using instrumental variables techniques, one still controls for the endogeneity of unmeasured characteristics. However, one does not run the risk of contaminating demand estimates using inappropriate supply-side assumptions. The approach proposed herein is amenable to being extended to include the supply side. However, our focus in the current paper is on the estimates of the brand choice model only.

<sup>1</sup> The simulation results are not included in the paper given space considerations, but they are available from the authors.

The main problem with accounting for the presence of UBCs is that we cannot treat them as just another error component, assume a distribution, and integrate them out of the demand function. If the UBCs are correlated with prices, then the latter would vary as we integrate over the support of the distribution of the former. Hence prices cannot be assumed to be fixed when integrating out the UBCs. Typical approaches to controlling for price endogeneity have therefore been computationally burdensome.

## 2.1. Maximum Likelihood Procedures

One approach is to model the joint distribution of prices and choices using maximum likelihood methods. One can then integrate out the UBCs and work with the unconditioned joint likelihood of prices and choices. Researchers have explored both full-information and limited-information approaches when modeling prices. In the former, prices are modeled structurally as the equilibrium outcome of a static Bertrand game between firms (Villas-Boas and Zhao 2001, Sudhir 2001, Yang et al. 2003). In the latter, prices are modeled in reduced form as a linear function of variables that are assumed to be uncorrelated with the UBCs (V-BW).<sup>2</sup>

In both full and limited information approaches, one estimates a system of equations that are nonlinear in some or all of the endogenous variables:

$$f(Q_t, p_t, X_t; \Theta) = u_t,$$

where  $Q_t$  is a vector of quantities sold for each product in week  $t$ ,  $p_t$  captures the corresponding retail prices in week  $t$ ,  $X_t$  captures exogenous demand and supply-shifting variables,  $u_t$  are UBCs and random supply (price) shocks drawn from a normal distribution in week  $t$ , and  $\Theta$  are the model parameters to be estimated. Because the endogenous variables contained in  $Q_t$  and  $p_t$  enter the model nonlinearly, one must use a transformation of variables to derive the joint likelihood of the data:

$$L = \sum_t \log \left\| \frac{\partial f_t}{\partial(Q_t, p_t)} \right\| - \frac{T}{2} \log \left| \frac{1}{T} \sum_t f_t f_t' \right|, \quad (1)$$

where the Jacobian term,  $\partial f_t / \partial(Q_t, p_t)$ , is required in the likelihood function to account for the transformation. When this term is overlooked, it leads to incorrect specifications of the likelihood function. Accounting for preference, heterogeneity would complicate the form of the joint likelihood function for choices and prices (either structurally or in reduced form) because heterogeneity has to be integrated out over each household's choice history that spans

several weeks, but the UBCs and random supply (price) shocks vary weekly. Without adequate controls for heterogeneity, one may not be able to distinguish whether parameter bias is associated with price endogeneity or omitted heterogeneity. To avoid the transformation of variables, Draganska and Jain (2002) use a simulation method that recovers the likelihood by using local averaging over draws of prices and choices from the equilibrium induced by the model. Although this approach does not fully account for heterogeneity, it nevertheless represents an innovative approach to dealing with the maximum-likelihood estimation (MLE)-based estimation problem.

Yang et al. (2003) propose a hierarchical Bayesian approach to estimating the parameters of the logit demand-based Bertrand equilibrium model, while accounting fully for heterogeneity. This study represents a methodological breakthrough in its ability to estimate a full information maximum likelihood demand and supply model. On the supply side, prices are assumed to be generated as the outcome of a Bertrand equilibrium:

$$p_{jt} = c_{jt} - \left( \sum_{k \in \text{product line}} \frac{\partial Q_{kt}}{\partial p_{kt}} \right)^{-1} Q_{jt} + v_{jt}, \quad (2)$$

where  $c_{jt} + v_{jt}$  is the marginal cost of brand  $j$  at time  $t$ ,  $v_{jt}$  is a normally distributed random component of costs, and  $(\sum_{k \in \text{product line}} \partial Q_{kt} / \partial p_{kt})^{-1} Q_{jt}$  is the equilibrium mark-up.<sup>3</sup> However, as reflected by the three published discussions following that paper, there are several unresolved issues with this approach. First, as we demonstrate in our Monte Carlo simulation study, the approach is vulnerable to misspecification in the mechanism that generates the correlation between prices and the UBCs. There are two potential sources of misspecification: (1) If manufacturers in the marketplace are not pricing according to the assumed Bertrand-Nash assumption, this will lead to biased and inconsistent estimates, not only of the parameters in the pricing equation, but more importantly of the demand parameters.<sup>4</sup> (2) The correlation could stem not from supply-side behavior, but from demand-side effects. For instance, if prices are persistent over time

<sup>3</sup>Note that the mark-up term contains all the price elasticities of demand of the products contained in the same product line as brand  $j$ . This is the outcome of the multiproduct pricing problem. In the case of a retailer setting the category profit-maximizing prices, the summation would be over all the products in the category.

<sup>4</sup>In practice, this model may not constitute a good representation of retail price variation. Retail prices may often exhibit large temporary discounts that are not reflected in wholesale prices. Erdem et al. (2003) find that retail prices in other consumer package goods (CPG) categories appear to follow a two-state Markov switching process: regular price versus temporary discount.

<sup>2</sup>V-BW use lagged prices (or shares) as well as production cost proxies.

and consumers stockpile, then the average level of omitted inventories in a given week may be correlated with prices. Erdem et al. (2003) resolve this type of problem by modeling consumers' dynamic purchase decisions and integrating the unobserved inventories out of the likelihood. Alternatively, prices or promotional variables may affect whether or not a product enters a consumer's "consideration set" (Mehta et al. 2003, Van Nierop et al. 2001). If consideration sets are omitted, one would obtain brand-specific error components that are correlated with prices (e.g., no consideration would be an error component with value negative infinite, and certain consideration would be an error component with value zero). In both cases, the previous research has documented biases in the same direction as those resulting from our UBCs.<sup>5</sup> Unlike Yang et al. (2003), our proposed approach can accommodate a variety of supply-side behavior, as well as UBCs stemming from demand-side effects.

A second concern with a full-information approach with heterogeneous demand is that even if the chosen pricing model (e.g., Bertrand-Nash) is correct, the equilibrium in prices may not be unique, especially in the case of multiproduct firms (e.g., the retailer). In the presence of multiple price equilibria, one cannot invoke the transformation-of-variables theorem and, hence, the likelihood function (1) cannot be derived. For a discussion of this and other trade-offs between full and limited information approaches, we refer the reader to the published discussions accompanying Yang et al. (2003).

The limited information maximum likelihood approach (V-BW) can be thought of as a reduced-form of the Equilibrium Model (2) above:

$$p_{jt} = c_{jt} + \omega_{jt}, \quad (3)$$

where  $c_{jt}$  is still the observed component of marginal costs for brand  $j$  in period  $t$ , and  $\omega_{jt}$  is a normally distributed composite error that contains both the firm's mark-up and the unobserved component of marginal costs. Because  $\omega_{jt}$  contains the mark-up, it will covary with demand factors including the UBCs and, hence, one must also specify a joint distribution for the random components of demand and supply. The limited information approach is, in principle, more flexible. However, the approach requires estimating covariances between the supply and demand shocks to resolve the endogeneity of prices. These covariance terms are identified off the

weekly covariation in mean utilities and shelf prices. There are two possible problems here. If the assumptions involving the joint distribution are incorrect, then we may obtain incorrect estimates for the model parameters. For example,  $\lambda$  might reflect unmeasured promotions, consumer inventories, advertising, or other usual explanations used to motivate its inclusion, all of which would likely exhibit substantial time dependence. Similarly, it is unlikely the reduced form of (2) would induce a Gaussian distribution on  $\omega$ . How well this method works in accounting for endogeneity depends crucially on the data required to pin down the covariances.

In the limited information spirit of V-BW, Petrin and Train (2002) suggest a computationally simple two-stage procedure termed the "control function" approach. In the first stage, they regress the endogenous variable, price, on instrumental variables that are correlated with price but uncorrelated with UBCs. In the second stage, the residual from the price regression is included as a brand covariate in the choice model to control for the UBCs. The manner in which the "control function," i.e., the functional form for including the residual in the demand function, enters the second-stage choice model determines whether one obtains the correct estimates of the model parameters. Assuming the wrong function of the residual leads to specification error.<sup>6</sup> The approach also assumes that all the information about the UBC is contained in the observed price variation. If, for instance, one expects that retail prices are sticky (costly adjustments lead firms to vary prices only when the benefits exceed the costs), then price variation alone may not contain all the information related to UBCs.

## 2.2. Instrumental Variables Procedures

A separate literature has suggested instrumental variables (IV) procedures to handle the price endogeneity problem. Although these IV approaches are potentially less efficient, they do not require additional structural assumptions about the supply side, hence IV approaches are more robust because they are less vulnerable to specification error. Nevertheless, the nonlinearity of the logit model prevents a straightforward application of standard linear IV methods to resolve endogeneity due to the presence of UBCs. Berry (1994) proposes an inversion procedure to obtain an econometric model that is linear in prices and, thus, facilitates the use of standard IV procedures. We refer the reader to Berry et al. (1995) and Nevo (2001) for applications with aggregate data.

<sup>5</sup>Note that the direction of bias relates to the sign of the correlation between prices and UBCs. If these were negatively correlated, then one would expect an upward bias in the mean price response parameter. In this instance, ignoring UBCs would overstate the sensitivity to prices. We thank an anonymous reviewer for pointing this out.

<sup>6</sup>The control function approach was originally proposed by Blundell and Powell (2001), who estimate control functions semi-parametrically in the case of binary choices.

Goolsbee and Petrin (2001) carry out the analogous exercise using individual choice data. Using cross-sectional household data across several geographic markets, they calibrate market-level brand intercepts that exactly match the predicted aggregate shares in each market to the observed market shares in their household sample. The mean price response is obtained by regressing these market-specific brand intercepts on market-specific prices and other brand-related variables, using instrumental variables to control for price endogeneity. Their approach is closest in spirit to the method we use in this paper to account for endogeneity. The principal difference is that Goolsbee and Petrin (2001) use cross-sectional data and hence do not deal with panel data issues such as unobserved heterogeneity and state dependence. Our approach, in contrast, provides maximum likelihood estimates of all the heterogeneity parameters.<sup>7</sup>

One issue with instrumental variables procedures is that, although they provide consistent estimates, they are not efficient. We will examine this issue in the context of our empirical results when we discuss various robustness checks associated with our analysis.

### 3. Model

We now discuss the model of consumer choice within a supermarket product category. First we discuss the standard random coefficients logit choice model. We then show how UBCs might enter this model and discuss why, even with household data, the UBCs might generate endogeneity bias.<sup>8</sup>

We follow the standard random utility approach (McFadden 1981, Guadagni and Little 1983). We assume that on a given shopping trip consumers either choose a single unit of the brand giving them the highest utility in the category, or choose not to purchase in the category. We also assume the utility derived from any alternative may be written as a function of intrinsic brand preferences, the effects of marketing conditions such as whether the item

is featured or displayed, and the effects of state dependence. Note that we do not model the decision of when to go to the store or which store to choose. We focus on purchase incidence and brand choice conditional on a store visit. The choice probabilities associated with each of the brands and the no-purchase “alternatives,” are given by the logit model. One could instead specify a nested logit structure on the purchase incidence decision. Because we allow brand preferences to be correlated, we do allow for (unconditional) heterogeneity in consumer switching between purchase and no-purchase. We do not expect the additional structure of the nested logit (e.g., the additional random coefficient) to change our main findings for the role of unmeasured brand characteristics.

Suppose there are  $J$  brands and the  $J + 1$  no-purchase option in the category. We describe a product  $j$  during week  $t$  by its intrinsic brand value, a  $(K \times 1)$  vector of observed (to the econometrician) marketing variables (other than price),  $X_{jt}$ , and retail shelf price,  $p_{jt}$ . Unlike aggregate data analyses, we also include household-specific variables such as state dependence, which are constructed using the household’s purchase history. These variables are denoted by  $Y_{hjt}$ , where  $h$  denotes the household. On a trip during week  $t$ , household  $h$ ’s conditional indirect utility for brand  $j$  is given by

$$u_{hjt} = \xi_{hj} + X'_{jt}\beta_h - \alpha_h p_{jt} + Y'_{hjt}\gamma + \varepsilon_{hjt}, \quad j = 1, \dots, J \quad (4)$$

$$u_{h(J+1)t} = \varepsilon_{h(J+1)t}, \quad \text{if no purchase,}$$

where  $\xi_{hj}$  is the brand intercept (intrinsic brand value),  $\beta_h$  is the vector of household  $h$ ’s tastes for characteristics,  $\alpha_h$  is the price sensitivity,  $\gamma$  is the effects of household-specific variables, and  $\varepsilon_{hjt}$  is an i.i.d. random taste shock drawn from a Type I extreme value distribution. We have also normalized the no-purchase alternative to give expected utility zero.<sup>9</sup>

To capture heterogeneity in consumer preferences and response parameters, we model the taste vector  $(\xi_h, \beta_h, \alpha_h)'$  as a random draw from a multivariate normal distribution  $N([\bar{\xi} \ \bar{\beta} \ \bar{\alpha}]', \Omega)$ :

$$\begin{bmatrix} \xi_h \\ \beta_h \\ \alpha_h \end{bmatrix} = \begin{bmatrix} \bar{\xi} \\ \bar{\beta} \\ \bar{\alpha} \end{bmatrix} + \Omega^{1/2}\eta_h, \quad \eta_h \sim N(0, I_{(J+K+1)}),$$

<sup>7</sup> Kuksov and Villas-Boas (2003) propose a similar IV estimator. Using the partial likelihood of the data, they construct moment conditions that are used to fit demand parameters while instrumenting for policy variables such as prices. This “quasilielihood” approach provides consistent, but less-efficient, estimates. This is because the procedure uses the densities of each observation, not the joint distribution.

<sup>8</sup> In our empirical application, we use a category in which consumers do make single unit purchases. More generally, for categories with quantity choices (i.e., consumers regularly purchase more than one unit) it would be straightforward to implement the IV procedure discussed below in the context of models that accommodate quantity choices (e.g., Chiang 1991) or multiple brand purchases (Dubé 2004, Kim et al. 2002). We thank an anonymous reviewer for suggesting these extensions.

<sup>9</sup> Often, the no-purchase decision is defined as choosing other alternatives in the category not captured by the  $J$ -modeled brands. Because our data contain information for all shopping trips, we define the no-purchase alternative as allocating the entire shopping budget to other categories.

where  $(\bar{\xi}, \bar{\beta}, \bar{\alpha})'$  are the mean parameters of the distribution of heterogeneity and  $\Omega^{1/2}\eta_h$  are household-specific deviations from the mean. The term  $\eta$  is a vector of standard normal deviates and  $\Omega^{1/2}$  is the lower-triangular Cholesky factor of  $\Omega$ . The vector  $(\bar{\xi}, \bar{\beta}, \bar{\alpha})'$  and the matrix  $\Omega^{1/2}$  consist of parameters to be estimated. Gathering terms, we can rewrite a household's conditional indirect utility for brand  $j$  at date  $t$  as

$$u_{hjt} = \bar{\xi}_j + X'_{jt}\bar{\beta} - \bar{\alpha}p_{jt} + Y'_{hjt}\gamma + [1, X'_{jt}, p_{jt}]\Omega^{1/2}\eta_h + \varepsilon_{hjt}. \quad (5)$$

Integrating out the extreme value error term,  $\varepsilon$ , the probability that a consumer chooses alternative  $j$  has the logit form

$$P_{hjt} = \exp(\bar{\xi}_j + X'_{jt}\bar{\beta} - \bar{\alpha}p_{jt} + Y'_{hjt}\gamma + [1, X'_{jt}, p_{jt}]\Omega^{1/2}\eta_h) \cdot \left[ 1 + \sum_{i=1}^J \exp(\bar{\xi}_i + X'_{it}\bar{\beta} - \bar{\alpha}p_{it} + Y'_{hit}\gamma + [1, X'_{it}, p_{it}]\Omega^{1/2}\eta_h) \right]^{-1}, \quad j = 1, \dots, J + 1. \quad (6)$$

This model has the usual random coefficients logit form, which can be estimated using MLE.

Now suppose that there are factors unobserved by the researcher that influence the households' choice probabilities. These factors can be of several kinds: Factors that vary across brands  $j$ , households  $h$ , and time  $t$  will be absorbed by the extreme value error term,  $\varepsilon_{hjt}$ . Those that are brand specific but are invariant across households and time will be part of the intrinsic brand preference term,  $\bar{\xi}$ . Factors that are only household specific but do not vary across brands  $j$  and time  $t$  will drop out of the logit expression above. Proceeding in this manner, what is not accounted by the above logit probability expression are characteristics that are common across households but are specific to the alternative and time, i.e.,  $jt$ . We refer to such factors as UBCs and denote them by  $\lambda_{jt}$ . Included in this are factors that vary over time and are common to all the alternatives except the outside good. In other words, they can include factors that are shared by all the brands within the category. The most likely source of such factors is the retail environment that is common to all households and has a similar effect on all households, but is not typically included in scanner panel datasets. The corresponding conditional probability that a consumer chooses alternative  $j$  has the form

$$P_{hjt} = \exp(\bar{\xi}_j + X'_{jt}\bar{\beta} - \bar{\alpha}p_{jt} + Y'_{hjt}\gamma + [1, X'_{jt}, p_{jt}]\Omega^{1/2}\eta_h + \lambda_{jt})$$

$$\cdot \left[ 1 + \sum_{i=1}^J \exp(\bar{\xi}_i + X'_{it}\bar{\beta} - \bar{\alpha}p_{it} + Y'_{hit}\gamma + [1, X'_{it}, p_{it}]\Omega^{1/2}\eta_h + \lambda_{it}) \right]^{-1}, \quad j = 1, \dots, J + 1. \quad (7)$$

We expect two potential problems to arise if one ignores the role of  $\lambda_{jt}$ . First, ignoring UBCs would force the model to absorb these effects in the heterogeneity distribution of preferences and sensitivities to marketing and other activities, i.e., in the random coefficients. Thus, one could misattribute differences in consumer behavior associated with varying  $\lambda_{jt}$  as heterogeneity. A second concern is related to the discussion in Berry (1994). If marketing managers set prices strategically, then one might expect the price-cost margins to be functions of the underlying product-related factors  $(\bar{\xi}_j, X_{jt}, \lambda_{jt})'$  for all  $j$ . The induced correlation between prices,  $p_{jt}$ , and unobserved attributes,  $\lambda_{jt}$ , may bias the estimates of the model parameters.

## 4. Empirical Approach

### 4.1. Estimation

We now discuss the estimation of the model parameters described in the previous section. Note that we first estimate the parameters characterizing the moments of the population distribution of tastes (the mean,  $(\bar{\xi}, \bar{\beta}, \bar{\alpha})'$ , and the variance,  $\Omega$ ). Conditional on the parameters characterizing the distribution of heterogeneity, we then obtain individual-level parameters via Bayes Rule.

We adapt the two-stage procedure suggested in the extensions of Berry (1994). To simplify the notation, we define  $\delta_{jt} = X'_{jt}\bar{\beta} - \bar{\alpha}p_{jt} + \bar{\xi}_j + \lambda_{jt}$ , to capture the mean utility of brand  $j$  across households, and  $\mu_{hjt} = Y'_{hjt}\gamma + (1, X'_{jt}, p_{jt})\Omega^{1/2}\eta_h$  to capture household-specific deviations from the mean. We can rewrite the conditional probability that consumer  $h$  chooses brand  $j$  as

$$P_{hjt} = \frac{\exp(\delta_{jt} + \mu_{hjt}(X_{jt}, p_{jt}, Y_{hjt}, \eta; \Omega^{1/2}, \gamma))}{1 + \sum_{i=1}^J \exp(\delta_{it} + \mu_{hit}(X_{it}, p_{it}, Y_{hit}, \eta; \Omega^{1/2}, \gamma))}, \quad j = 1, \dots, J + 1. \quad (8)$$

If one were estimating a standard random coefficients logit model, then the mean parameters  $\bar{\alpha}$ ,  $\bar{\beta}$ , and  $\bar{\xi}$  would be jointly estimated, along with the parameters characterizing the distribution of heterogeneity using MLE. However, in the current context, the "parameters of interest" in Equation (8) above are  $\delta_{jt}$  in addition to the parameters of the heterogeneity distribution and  $\gamma$ . Note that these  $\delta_{jt}$  contain not only the mean parameters  $\bar{\alpha}$ ,  $\bar{\beta}$ , and  $\bar{\xi}$ , but also the

UBCs  $\lambda_{jt}$ . So, by estimating  $\delta_{jt}$  as time-varying brand intercepts, we are capturing the net effect of all these factors on choice probabilities. As with the standard random coefficients model, the unknown parameters  $(\{\delta_{jt}\}_{j,t}, \gamma, \Omega^{1/2})$  are estimated by maximizing the corresponding sample likelihood function. In the second stage, we project the estimated  $\hat{\delta}_{jt}$  onto the brand intercepts and the marketing variables to recover the mean parameters of the distribution of heterogeneity, i.e.  $\bar{\alpha}$ ,  $\bar{\beta}$ , and  $\bar{\xi}$ . In this second stage, we do two things. First, we use instrumental variables to control for the endogeneity of prices. Second, we account for the estimation error in  $\delta_{jt}$  from the first stage. Note that the residual in the second stage is nothing but the UBC. By looking at the variance in  $\delta_{jt}$  explained by the brand intercepts and marketing variables, we can determine the relative importance of observed versus unobserved brand characteristics. Note also that we do not impose any additional model structure to control for potential UBCs. If the UBCs do not play an important role in governing individual choices, then in the second stage we should recover taste parameters identical to those of the usual random coefficients logit model.

We can now write down the likelihood of a household  $h$ 's purchase history:

$$L_h(\{\delta_{jt}\}_{j=1, t=1}^{T, J}, \Omega^{1/2}, \gamma) = \int \prod_{t=1}^{T_h} \prod_{j=1}^{J+1} P_{hjt}^{y_{hjt}} \phi(\eta) \partial \eta, \quad (9)$$

where

$$y_{hjt} = \begin{cases} 1, & \text{if alternative } j \text{ is selected} \\ 0, & \text{else} \end{cases}$$

and  $\phi(\cdot)$  is the probability density function for a standard normal. We are now ready to carry out the first stage, during which we find values of  $\delta_{jt}$  and  $\Omega^{1/2}$  to maximize the log-likelihood:

$$l(\{\delta_{jt}\}_{j=1, t=1}^{T, J}, \Omega^{1/2}) = \sum_{h=1}^H \log(L_h). \quad (10)$$

Because the computation of (9) involves evaluating a complex multivariate integral, we use Monte Carlo simulation methods, or “simulated maximum likelihood” (see Keane 1993, and Hajivassiliou and McFadden 1993). For each household, we take  $R$  draws of  $\eta_h$  from the distribution  $\phi(\cdot)$  and simulate the likelihood using

$$\hat{L}_h = \frac{1}{R} \sum_{r=1}^R \prod_{t=1}^{T_h} \prod_{j=1}^{J+1} P_{hjt}^{y_{hjt}}(\eta_{hr}). \quad (11)$$

In order to recover the mean taste parameters  $(\bar{\xi}, \bar{\beta}, \bar{\alpha})'$ , we project the estimated time-varying intercepts  $\hat{\delta}_{jt}$  onto the product attributes and prices:

$$\hat{\delta}_{jt} = X_{jt}\beta - \alpha p_{jt} + \xi_j + \lambda_{jt}. \quad (12)$$

Because we expect prices to be correlated with the remaining unobserved (to the econometrician) brand characteristics,  $\lambda_{jt}$ , we use  $N$  exogenous variables  $Z_{jt}$  to instrument for prices. We stack the  $\hat{\delta}_{jt}$  into the  $(JT \times 1)$  vector  $\hat{\delta}$ , the instruments into the  $(JT \times N)$  matrix  $Z$ , the covariates  $(1_j, p_{jt}, X_{jt})'$  into the  $(JT \times (K+2))$  matrix  $X$ , and the UBCs into  $(JT \times 1)$  vector  $\lambda$ . We also assume  $E(Z'\lambda) = 0$ . To control for the uncertainty in using the estimated  $\hat{\delta}_{jt}$ , we use a minimum distance procedure that takes the covariance matrix of these estimates,  $\hat{\vartheta} = \widehat{\text{Var}}(\hat{\delta})$ , as the distance matrix. Because the covariance matrix is known we use generalized two-stage least squares:

$$\begin{bmatrix} \hat{\xi} \\ \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = (X'P_z\hat{\vartheta}^{-1}P_zX)^{-1}X'P_z\hat{\vartheta}^{-1}\hat{\delta},$$

where  $P_z = Z(Z'Z)^{-1}Z'$  is the projection operator for  $Z$ . Note that if in fact UBCs do not account for a notable proportion of the variation in  $\delta_{jt}$  across weeks and brands, then the estimates obtained from the two-stage procedure would be roughly identical to the standard random coefficients logit model. That is, we should empirically recover the same heterogeneity parameters and the same mean responses to marketing variables.

An advantage of the proposed approach is that we do not need to make any specific parametric assumptions about the price-generation process. Thus, we do not need to assume a specific retail or channel model for the determination of prices. Moreover, we do not need to specify a parametric distribution for the unobserved attributes  $\lambda_{jt}$ . Finally, we do not need to evaluate the joint-likelihood of prices and choices. Although we use maximum likelihood methods to estimate the model parameters, we are able to avoid the computational problems discussed in an earlier section. A limitation of our approach is the need to estimate time-specific brand intercepts. In scenarios for which one has too many brands to handle the computational burden of estimating the full set of intercepts, one may need to resort to the numerical inversion procedure of Goolsbee and Petrin (2001). Note that because they use cross-sectional data, one would need to modify their inversion procedure to reflect the panel structure of scanner data. We also propose using instrumental variables, which will lead to a loss in efficiency because one loses some of the observed price variation when projecting first onto exogenous instruments. Finally, the two-stage approach identifies the mean taste parameters from the variation in aggregate weekly data, rather than the individual choice data, which could also lead to a loss in precision. We explore these trade-offs in the results section, §6.

### 4.2. Targeting

After completing the two-stage procedure, we use the estimated mean parameters,  $\bar{\Theta} = (\bar{\beta}, \bar{\alpha}, \bar{\xi})'$ , and variance,  $\Omega$ , to characterize the population distribution of tastes. We apply Bayes Rule to calculate an individual household's expected preference parameters conditional on the household's purchase history and the population distribution of tastes:

$$\hat{\Theta}_h = \int \frac{\Theta L_h(\eta | \bar{\Theta}, \Omega) \phi(\eta) \partial \eta}{L_h(\eta | \bar{\Theta}, \Omega) \phi(\eta) \partial \eta}. \quad (13)$$

Note that this approach, termed "approximate Bayesian" (Allenby and Rossi 1999), neither directly accounts for the uncertainty in household parameters due to estimation error in the population taste parameters, nor the variation in individual choice histories. Revelt and Train (1999) propose a bootstrap procedure to solve the former problem by resampling the parameters of the population distribution of tastes from their joint asymptotic distribution and computing the corresponding conditional expectations in (13) for each draw.<sup>10</sup> This correction is important in the current context, because our two-step estimator may have lower precision than Direct MLE. The second form of uncertainty is due to the variation within a household's observed choice history: Some households may have more informative histories than others. We do not address this concern currently, because it affects Direct MLE and the two-step estimators in the same way. However, one could easily resolve this additional source of uncertainty by calculating each household's conditional variance in tastes along with the conditional means.

In addition to characterizing the distribution of consumer tastes, marketers may also wish to use the results to design targeted marketing policies, such as first-degree price discrimination. In the context of the scanner data used in the next section, targeted prices should enable the retailer to improve profits compared with the uniform pricing policy that was used to generate the observed shelf prices (Rossi et al. 1996). For a given set of prices,  $\mathbf{p}$ , Bayes Rule provides household  $h$ 's expected probability of purchasing alternative  $j$  conditional on its purchase history:

$$P_{hj}(\mathbf{p}) = \int \frac{P_{hj}(\mathbf{p}, \eta | \bar{\beta}, \Omega^{1/2}) L_h(\eta | \bar{\beta}, \Omega^{1/2}) \phi(\eta) \partial \eta}{L_h(\eta | \bar{\beta}, \Omega^{1/2}) \phi(\eta) \partial \eta}, \quad j = 1, \dots, J. \quad (14)$$

To compute optimal targeted prices, the retailer solves

$$\max_{p_1, \dots, p_J} \sum_{j=1}^J (p_j - w_j) P_{hj}(\mathbf{p}), \quad (15)$$

<sup>10</sup> The asymptotic properties of the random coefficients logit model are derived in McFadden and Train (2000).

where  $w_j$  is the wholesale price of product  $j$ . As above, one could use a similar bootstrap procedure to account for the uncertainty in the MLE parameters and individual histories. The category profits depend on the retail prices to be computed, wholesale prices assumed to be known, and the household's expected probability of purchasing alternative  $j$  conditional on its purchase history.

In the empirical section below, we compare the posterior distribution of consumer preference, the response parameters, and the targeted prices under three model specifications. The three models compared are the two-step IV/MLE model that uses instrumental variables to account for endogeneity in the second stage, the two-step OLS/MLE model that allows for the presence of UBCs but does not account for price endogeneity in the second stage (uses GLS), and the Direct MLE model where no UBCs are included.

### 4.3. Simulation Results

In an appendix (available at <http://mansci.pubs.informs.org/ecompanion.html>), we compare the theoretical differences between limited- and full-information approaches relative to our proposed approach. We then carry out several Monte Carlo simulations to compare the empirical properties of the three approaches. The main goal of the simulations is to illustrate some of the limitations of using more assumed structure as in the limited- and full-information maximum likelihood approaches discussed in §2.1, hereafter LIML and FIML respectively. In general, we find our proposed approach provides robust results across several alternative mechanisms that generate the endogeneity problem. In contrast, FIML and LIML approaches tend to be more sensitive to misspecification problems. Additionally, we run simulations with random coefficients models that demonstrate the presence of both forms of bias—in the mean, and the variance of the distribution of heterogeneity when endogeneity is ignored. Although not verified in the simulations, insofar as LIML may not resolve endogeneity with typical short scanner panels (see below), one might be concerned that this approach could yield overly disperse predictions for the distribution of heterogeneity. We briefly describe some of our findings below. We refer the reader to the online appendix to obtain specific details on the construction of the simulations and the findings.

To assess the robustness of our approach, we simulate data for a number of price-generation mechanisms. For each simulated data type, we compare estimates using the static LIML and FIML approaches (see §2.1) with our proposed two-step approach (see §4.1). We also estimate the standard conditional logit (i.e., no account for the supply side) as a benchmark. In the first four cases, we generate data from



variations of the LIML model. That is, we generate data from a linear price specification with an additive error term. In each of these four cases, the FIML approach is misspecified because it incorrectly assumes prices are a nonlinear function of demand parameters (the mark-up term). In the first simulation, we use an autocorrelated supply shock that is contemporaneously correlated with the demand shocks. As expected, the conditional logit does not recover the true parameters. Surprisingly, the FIML approach is better able to recover the parameters than the LIML approach, probably because it is a more parsimonious specification. Finally, the proposed approach is best able to recover the true parameters. Naturally, this last result is predicated on the assumption that the instruments used are valid (i.e., they are uncorrelated with the demand and supply shocks). In the second simulation, we retain the autocorrelated supply shock, but eliminate the correlation between supply and demand shocks. In this case, the results are comparable to the first case.

In the third simulation, we assume that the supply shocks are independent over time, but are correlated with demand shocks. In this case, the LIML model is now correctly specified and, as expected, performs very well. Nevertheless, the proposed approach appears to provide a comparable level of accuracy in the results. In the fourth simulation, we assume supply and demand shocks are independent. As in the previous case, both the proposed approach and LIML perform well. However, FIML now performs very poorly because it incorrectly makes the implicit assumption of correlation between prices and demand shocks due to the mark-up term.

In the fifth simulation, we explore the result in Erdem et al. (2003) whereby observed prices are found to follow a Markov process. We generate prices from a Markov process to capture this result. In this case, both the LIML and FIML approaches would be misspecified, because both assume an incorrect density of prices. Despite the misspecification, the FIML results are still reasonable, although the proposed approach is noticeably more accurate. This result likely occurs because the FIML model is an implicit function of the prices, via the mark-up, and hence is able to capture some of the Markov-switching behavior.

In the final two simulations, we generate data from variations of the FIML model. In both cases, the proposed approach fares very well, whereas the other approaches have mixed results. In the sixth simulation, we use the Bertrand model and we allow for additional correlation between the supply and the demand shocks. Both the logit and the LIML approach perform quite poorly. The FIML approach gets much closer to the true parameters, but there is still bias. This latter result is due to the fact that

the endogeneity is only accounted for through the mark-up term. In the seventh simulation, we use the Bertrand model, but we assume the supply and demand shocks are uncorrelated. In this case, the FIML approach provides the most accurate results, as expected.

## 5. Empirical Results

### 5.1. Data

Our data were collected by ACNielsen for the Denver area between January 1993 and March 1995, a total of 117 weeks. On the demand side, we have the purchase histories of 2,100 households across all stores using checkout-counter scanners in the Denver Scantrak. In the current analysis, we use the purchase information for the margarine category. We focus on the 16-ounce size for the four largest brands, accounting for over 50% of the category volume.<sup>11</sup> Note that comparable results are available for the refrigerated orange juice category, on request. We focus on consumer purchases in the largest chain, where we have the corresponding weekly prices and marketing mix variables for all the alternatives that are available in the category. For estimation in a particular category, we only use those households that purchase a brand in the category at least once. This criterion leaves us with 992 households for margarine. Descriptive statistics for the purchase data are provided in Table 1. An attractive feature of our data are the long household purchase histories. On average, households purchasing margarine make roughly seventy three trips over the 117 weeks. Similarly, the households purchase margarine six times during this period, on average.<sup>12</sup> Although we have fairly long histories for each household, six purchase occasions is probably insufficient to estimate a simple logit household-by-household. We also create a loyalty variable consisting of an indicator for whether or not the brand chosen by a household on a given trip was also chosen on the previous trip during which a purchase occurred.

In the second stage, we project the weekly brand intercepts onto the weekly prices, feature ad, and display conditions. The Nielsen data also include these marketing variables for all the alternatives each

<sup>11</sup> Researchers frequently drop “fringe” brands with very small market shares. It is possible that the average price level of the fringe segment could correlate with demand for national brands. This correlation could also be a potential source for the UBC biases we discuss.

<sup>12</sup> We drop weeks during which any single alternative is chosen by fewer than three households. Thus, we drop 12 weeks in the margarine category. We estimate the typical conditional logit with and without these weeks and find no change in our parameters. Hence, we do not expect that dropping these weeks will impact our main findings.

**Table 1** Descriptive Statistics for Margarine Shopping Trips

	Mean	Std	Min	Max
Choose Blue Bonnet (conditional on purchase)	0.31	0.46	0	1
Choose I Can't Believe It's Not Butter (conditional on purchase)	0.27	0.44	0	1
Choose Parkay (conditional on purchase)	0.27	0.45	0	1
Choose Shedd's (conditional on purchase)	0.14	0.35	0	1
Trips per household	72.9	57.73	1	470
Purchase trips per household	6.08	7.61	1	70
Total purchase trips		5,693		
Total trips		56,138		
Total households		992		

week corresponding to our purchase information. To control for the endogeneity of prices, we supplement our scanner data with instruments reflecting manufacturer-related variables. Previous work with store-level data has used such instruments as monthly factor price indexes collected by the Bureau of Labor Statistics (BLS). Even after filtering the monthly data into weekly time series, the instruments explain a very small portion of the observed price variation. We use weekly wholesale prices collected by Promodata Leemis Services. We expect these instruments to be highly correlated with shelf prices. However, because we do not expect the prices in a single store to affect the market wholesale price, we assume the instruments are uncorrelated with unmeasured product attributes within a given store-week. This assumption is analogous to the exogeneity assumptions used in previous work in this area using household- and store-level scanner data. For margarine, we also use the monthly BLS wholesale price index. These additional instruments, which are more similar to the typical form of instruments used, are intended to proxy for production costs associated with the products. In the section on robustness checks, we check for the exogeneity of our wholesale price series.

Descriptive statistics for the data used in the second-stage regressions appear in Table 2. In the margarine category, the brand "I Can't Believe It's Not Butter" (IC) has much higher prices than the competing alternatives. Similarly, IC has the highest wholesale prices. In contrast, Blue Bonnet (BB), the lowest-priced alternative, has the highest variance in both shelf and wholesale prices across weeks. BB also uses feature ads and in-aisle displays much more frequently than its competitors.

We now turn to the candidate instruments for shelf prices. In Table 3, we summarize the explanatory power of our various instruments for prices. We use  $R^2$  to capture the fraction of the price variation explained by a given instrument. Typically, as noted above, marketers have used factor prices to

**Table 2** In-Store Marketing Conditions for Margarine

	Mean	Std	Min	Max
Blue Bonnet price (cents per oz)	0.0361	0.0060	0.025	0.051
Blue Bonnet wholesale price (cents per oz)	0.0317	0.0032	0.028	0.036
Blue Bonnet feature ad	0.1804	0.3215	0	1
Blue Bonnet display	0.0329	0.0771	0	1
I Can't Believe It's Not Butter price (cents per oz)	0.0967	0.0044	0.087	0.103
I Can't Believe It's Not Butter wholesale price (cents per oz)	0.0654	0.0013	0.064	0.068
I Can't Believe It's Not Butter feature ad	0.0690	0.1726	0	1
I Can't Believe It's Not Butter display	0.0005	0.0052	0	1
Parkay price (cents per oz)	0.0518	0.0053	0.042	0.066
Parkay wholesale price (cents per oz)	0.0379	0.0018	0.034	0.041
Parkay feature ad	0.1438	0.2709	0	1
Parkay display	0.0218	0.0613	0	1
Shedd's price (cents per oz)	0.0604	0.0039	0.051	0.069
Shedd's wholesale price (cents per oz)	0.0392	0.0013	0.038	0.041
Shedd's feature ad	0.0286	0.1452	0	1
Shedd's display	0.0003	0.0029	0	1
Weeks		117		

proxy for the production costs associated with the various brands. When appropriate, others have used spot market commodity prices for raw ingredients, such as tomatoes for ketchup. However, these types of proxies for factor prices are not collected at the brand level and, thus, have no information about week-specific variation in prices across brands. For instance, occasional deep price discounts would likely not be recovered by these instruments. To illustrate the limitations of such instruments, we compare the explanatory power of brand-specific wholesale prices and factor price indexes (for margarine). As illustrated in Table 3, the wholesale price data we use explain a greater proportion of the variation in shelf prices than the instruments that do not vary across brands (BLS price index). In fact, wholesale prices alone explain nearly 80% of margarine price variation. In contrast, the product cost-related instruments capture less than 5% of the shelf price variation. For the category we use, the factor price proxies will clearly be inadequate for instrumenting prices, because we will lose most of the underlying price variation. In the final row of the table, we report the first-stage  $R^2$ , which represents a regression of prices on all exogenous variables, including the instruments. The

**Table 3** Analysis of Variance of Shelf Prices and the Explanatory Power of Instruments

	Margarine
Wholesale price	0.80
BLS	0.01
1st stage	0.97

corresponding  $F$ -statistic, 848.9, rejects the hypothesis that our instruments do not explain observed prices. This evidence illustrates how wholesale prices may be preferable instruments in terms of their relevance or their ability to explain price variation. However, if the unobserved quality is a market-level shock, then wholesale prices may be correlated with the terms  $\lambda_{jt}$ , and thus may not be valid instruments. We explore this issue in the robustness checks below.

## 5.2. Results

In this section, we present our results for the margarine category described above. To illustrate the importance of controlling for both endogeneity and heterogeneity, we present results for six sets of models. We have three scenarios to consider: direct estimation (standard brand choice model with non-time-varying brand intercepts), weekly brand fixed effects but no instruments, and weekly brand fixed effects and instruments. Each of these three scenarios is estimated with and without controls for unobserved consumer heterogeneity, leading to six sets of models. In scenarios two and three, only the mean taste coefficients obtained in the second stage will be different due to the use of instrumental variables, but the first-stage estimates will be identical. We do not instrument the feature and display variables.<sup>13</sup> We estimate each of these three models first using the homogeneous logit specification, and then using the heterogeneous, random-coefficients, logit specification. Note that, for the heterogeneous models, we do not report the covariance terms for our random effects to conserve space, the matrix  $\Omega = \Omega^{1/2}\Omega^{1/2}$  above. However, we allow for correlation in all the preference parameters to provide flexible substitution patterns. For consistency with the literature, we provide log-likelihood values, Akaike's information criterion (AIC) and Schwartz's Bayesian information criterion (BIC) to assess model fit. The likelihood and AIC both improve with the addition of heterogeneity and fixed effects. However, the BIC worsens with the addition of fixed effects because it penalizes additional parameters much more heavily than does the AIC. One should note that the BIC does not capture the fact that we are really only interested in the second-stage estimates of mean marketing response parameters after accounting for endogeneity. Below,

<sup>13</sup> We found that Hausman tests systematically failed to reject the hypothesis that these variables are exogenous. The Hausman test statistic is simply  $(\hat{\beta} - \hat{\beta}^{IV})\widehat{V}(\hat{\beta} - \hat{\beta}^{IV})$ , where  $\hat{\beta}$  is the uninstrumented parameter estimate,  $\hat{\beta}^{IV}$  is the instrumented estimate, and  $\widehat{V}$  is the estimated covariance of the difference between the two. This statistic is distributed chi-square under the null hypothesis that  $\beta = \beta^{IV}$ . Related findings have also been reported by Kuksov and Villas-Boas (2003).

we also look at hold-out predictions to assess relative model fit.

For the margarine category in Table 4, the first two columns illustrate that we obtain comparable estimates of the mean taste coefficients when we use the two-stage procedure versus the standard direct one-stage estimation approach. This result reassures us that we have sufficient information to estimate the weekly brand fixed effects with reasonable precision. To conserve space, we do not report the vector of weekly brand intercepts. However, all of these parameters are significant with  $t$ -statistics ranging from  $-4$  to  $-13$ . In the third column, we estimate the mean taste coefficients using instrumental variables for prices. We find that the point estimate of the mean price response increases dramatically after instrumenting. The high first-stage  $R^2$  is very encouraging. Our instruments explain roughly 97% of the variation in margarine prices. In contrast, other studies using proxies for production costs have reported  $R^2$  much lower values around 0.4. We now consider the effects of controlling for heterogeneity by comparing Columns 1 to 3 with Columns 4 to 6. As expected, we find that ignoring heterogeneity leads to a lower mean price response. As in Keane (1997), we also find that, although loyalty falls substantially after controlling for heterogeneity, it still remains significant. Moreover, we find loyalty remains significant even after controlling for endogeneity (two-step approaches). We also find sizeable increases in the price coefficient after instrumenting for price. A Hausman test for price endogeneity (two-step OLS/MLE versus two-step IV/MLE) yields a statistic of 26.34, rejecting the null hypothesis  $H_0: \beta^{OLS} = \beta^{IV}$ . The increase in magnitude of the mean price parameter is consistent with the previous findings in Besanko et al. (1998). The evidence suggests that ignoring both heterogeneity and endogeneity generates a downward bias in the parameter estimates.<sup>14</sup>

In comparing Columns 4 and 5, we obtain our key empirical finding that extends the current results on the endogeneity bias. We find that controlling for weekly brand-specific shocks leads to lower predicted dispersion in tastes, as captured by lower variances in the random coefficients. To see this point, we report the mean and standard deviation of each random coefficient in Table 5. For the price response, in particular, we see a sharp decline in the estimated standard deviation across households. This finding suggests

<sup>14</sup> Although not reported, removing the loyalty parameter from the model only impacts the estimated price coefficient. Dropping loyalty, the estimated price coefficient is smaller for all three cases (Direct MLE, two-step IV/MLE, two-step OLS/MLE). However, the magnitude of the bias is roughly the same.

**Table 4** Parameter Estimates for Margarine Demand

Covariate	Homogeneous						Heterogeneous							
	Direct MLE		Two-step OLS/MLE		Two-step IV/MLE		Direct MLE		Two-step OLS/MLE		Two-step IV/MLE		Two-step IV/MLE	
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
BB	-2.11	0.08	-2.08	0.11	-0.86	0.25	-1.82	0.07	-1.22	0.26	-0.72	0.26	-0.25	0.28
IC	1.48	0.18	1.39	0.30	5.05	0.66	1.74	0.14	3.23	0.63	4.57	0.66	4.96	0.71
PA	-1.21	0.10	-1.19	0.16	0.63	0.36	-0.58	0.07	0.80	0.34	1.25	0.35	1.32	0.38
SH	-1.08	0.12	-1.11	0.19	1.11	0.41	-0.77	0.10	0.20	0.40	0.98	0.42	1.22	0.44
Price	-58.28	1.79	-56.15	3.14	-95.55	6.77	-65.42	1.21	-83.40	6.47	-97.09	6.80	-93.84	7.23
Feature	0.32	0.05	0.29	0.05	0.16	0.06	0.20	0.06	0.06	0.10	0.11	0.10	-0.20	0.10
Display	1.01	0.20	0.92	0.24	0.60	0.27	1.04	0.31	1.42	0.40	1.84	0.39	0.43	0.41
Loyalty	1.64	0.03	1.66	0.03	1.66	0.03	0.59	0.04	0.55	0.06	0.55	0.06	0.55	0.06
Pseudo $R^2$			0.34		0.41				0.31		0.41		0.41	
First-stage $R^2$					0.97						0.97		0.97	
Log-likelihood	-24,071				-23,640		-23,535				-22,907			
AIC	48,158				48,122		47,141				46,713			
BIC	48,230				51,884		47,463				50,725			
Total trips							56,138							
Households							992							

Note. BB, Blue Bonnet; IC, I Can't Believe It's Not Butter; PA, Parkay; SH, Shedd's; AIC, Akaike's information criterion; BIC, Bayesian information criterion; coef., coefficient; s.e., standard error.

that ignoring unobserved time-varying brand characteristics, regardless of whether they generate endogeneity in prices, may lead to inflated estimates of heterogeneity. Hence, controlling for heterogeneity across brands may play a comparable role to heterogeneity in consumer tastes in terms of the identification of mean taste parameters. These results will have ramifications when used for marketing applications, such as targeting. The evidence suggests that failure to account for time-varying unobserved brand characteristics would cause a manager to overstate the degree of heterogeneity in a market. If the model were used for targeting, this could mislead a manager into expecting larger returns from price-discrimination policies. At the same time, managers have biased estimates of the mean response to marketing variables.

**Table 5** Means and Standard Deviations of Random Coefficients (Margarine)

Covariate	Direct MLE		Two-step OLS/MLE		Two-step IV/MLE	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
BB	-1.82	2.70	-1.22	1.52	-0.72	1.52
IC	1.74	3.51	3.23	2.68	4.57	2.68
PA	-0.58	2.88	0.80	1.22	1.25	1.22
SH	-0.77	2.72	0.20	2.39	0.98	2.39
Price	-65.42	45.71	-83.40	25.08	-97.09	25.08
Feature	0.20	0.06	0.06	0.08	0.11	0.08
Display	1.04	0.53	1.42	0.12	1.84	0.12

Note. BB, Blue Bonnet; IC, I Can't Believe It's Not Butter; PA, Parkay; SH, Shedd's.

In the seventh column of Table 4, we reestimate the two-step IV/MLE model with alternative instruments. If wholesale prices are correlated with the UBCs, then they would not be valid instruments. Note that if the correlation were positive, then in fact the size of the bias we report would be understated. Nevertheless, we use BLS wages for the food industry collected from the city market in which each brand is manufactured. For instance, because BB is produced in Minneapolis, we use wages in Minneapolis as an instrument for BB shelf prices. In general, these instruments are inferior to our wholesale prices in terms explaining the variation in shelf prices, although they still capture some of the brand-specific variation in prices. Although these instruments are not as informative about price variation as wholesale prices, they are unlikely to be correlated with UBCs. We find that the results do not change substantively, making us more confident our findings are not driven by inappropriate instruments. We do note that standard errors rise slightly. Hausman tests failed to reject the hypothesis that the price response parameter in Column 7 is statistically the same as that in Column 6.<sup>15</sup>

<sup>15</sup> The traditional Hausman test compares the OLS estimate (which is efficient and consistent under the null) with the IV estimate, which is consistent and inefficient under the null. The analogous test is slightly more complicated when comparing two different IV estimates because, under the null hypothesis, both estimates are inefficient. We use a parametric bootstrap to calculate the Hausman statistic. We use 5,000 independent draws from the asymptotic distribution of each IV estimate and calculate the mean and variance

**Table 6 Mean Own and Cross-Price Elasticities (Margarine)**

	Direct MLE					Two-step OLS/MLE					Two-step IV/MLE				
	BB	IC	PA	SH	No Purchase	BB	IC	PA	SH	No Purchase	BB	IC	PA	SH	No Purchase
BB	-1.97	0.04	0.13	0.07	0.07	-2.46	0.05	0.17	0.09	0.09	-2.86	0.06	0.20	0.11	0.11
IC	0.11	-5.25	0.12	0.21	0.20	0.15	-6.88	0.14	0.13	0.24	0.17	-8.00	0.16	0.15	0.27
PA	0.18	0.06	-3.13	0.09	0.10	0.20	0.06	-3.97	0.09	0.10	0.23	0.07	-4.62	0.11	0.12
SH	0.06	0.06	0.06	-3.77	0.06	0.07	0.04	0.06	-4.76	0.07	0.08	0.04	0.07	-5.55	0.08

Note: BB, Blue Bonnet; IC, I Can't Believe It's Not Butter; PA, Parkay; SH, Shedd's.

**5.3. Robustness Checks and Other Issues**

Although not reported, we also experimented with several other specifications to check the robustness of our findings. The objective was to consider other specifications that might account for the correlation between the error components and prices. We experimented with including demographic variables interacted with prices, with a nonlinear Box-Cox price response, with seasonality, and with promotional interaction terms. None of these changed our results substantively, so we do not report them, although they are available on request.

A second concern with our proposed approach is the inherent trade-off between bias, in the case of Direct MLE, and efficiency, in the case of our two-step approaches. The efficiency loss in the two-step approaches stems from both the fixed-effects procedure and the use of instrumental variables. By using fixed effects, the mean consumer taste parameters are identified by the weekly data, rather than the household-trip data. For instance, the standard error of the mean price response parameter increases from 1.2, in the case of Direct MLE, to 6.5, in the case of two-step OLS/MLE. In the current context, the decrease in precision does not seem problematic at face value, given the high *t*-statistic (*t* = 12.9 for two-step OLS/MLE). Similarly, we do not expect the instruments to be problematic in the current context because the *R*<sup>2</sup> for the first-stage regression of prices on all the exogenous instruments is 0.97. Comparing the standard errors for the mean price response under OLS/MLE versus IV/MLE, we observe a very small increase from 6.5 to 6.8.

Nonetheless, in general one may not have access to good instruments, which could lead to very imprecise estimates of the mean taste parameters. In such contexts, one would need to be careful to weigh the trade-off between the biased but efficient Direct MLE estimates, and a noisier IV estimate. We explore this matter further below in §5.5 by comparing the precision in household-level conditional mean taste parameters based on Direct MLE and the two-step approaches. A related problem could arise if one had access to

seemingly good instruments (i.e., highly correlated with prices) that are invalid (i.e., correlated with  $\lambda$ ). In this case, not only would the estimates be biased, they would also be less precise than the Direct MLE results.<sup>16</sup>

**5.4. Market Structure**

In the previous discussion, we have demonstrated the effects of ignoring either heterogeneity or endogeneity in logit brand choice models. The question, then, is are there situations in which we do not really need to account for UBCs? In other words, are the implications that we obtain from a model that does account for them similar to the implications that we obtain from a model without them? To answer this question, we provide the estimates for the mean own and cross-price elasticities for the margarine data in Table 6. Consistent with our earlier results, we find that the own-price elasticities are ordered (in magnitude) as follows: Direct MLE < two-step OLS/MLE < two-step IV/MLE. Hence, if one is interested in the magnitudes of the own-price elasticities, then one needs to be concerned about the effects of UBCs.

Now suppose the marketing manager is only interested in knowing the brands that are more or less price sensitive. In this case, all that matters are the relative magnitudes of the mean elasticities across brands. These relative elasticities are in Table 7. We find the relative magnitudes of the mean price elasticities are virtually identical across the three model specifications. In other words, if the focus is on understanding whether and by how much BB is more or less price elastic than Shedd's (SH) then the standard random coefficients model that allows for heterogeneity across consumers would be appropriate. In the absence of heterogeneity, this result can be shown theoretically because, for the homogeneous logit, the ratio of own-price elasticities for two brands *i* and *j* is independent of the price parameter:

$$\frac{\epsilon_{ii}}{\epsilon_{jj}} = \frac{(1 - s_i)p_i}{(1 - s_j)p_j}$$

of the corresponding differences in mean price parameters. These moments are then used to construct the statistic.

<sup>16</sup> Diagnosing this problem should be evident because estimates under each approach would be statistically indistinguishable. At the same time, the IV estimates would be much noisier.

**Table 7** Relative Own-Price Elasticity (Margarine)

	Ratio of own-price elasticities		
	MLE	OLS/MLE	IV/MLE
BB/IC	0.38	0.36	0.36
BB/PA	0.63	0.62	0.62
BB/SH	0.52	0.52	0.52
IC/PA	1.68	1.73	1.73
IC/SH	1.39	1.44	1.44
PA/SH	0.83	0.83	0.83

*Note.* BB, Blue Bonnet; IC, I Can't Believe It's Not Butter; PA, Parkay; SH, Shedd's.

where  $s_i$  and  $s_j$  are the market shares of brand  $i$  and  $j$  respectively. Although not reported, we find this result to hold in the context of the simulations described in the online appendix. It is when one is also interested in the point estimate of the mean price elasticity that the two-step IV/MLE model estimates need to be used.

Although not reported, we also examine the implications for market structure obtained from the two specifications. For the margarine category we estimate brand maps (as in Elrod and Keane 1995) by decomposing the covariance matrix of the heterogeneity distribution of brand preferences. We take the submatrix of  $\Omega$  corresponding to the intrinsic brand preferences,  $\Omega_\xi$ . Taking the Cholesky decomposition of  $\Omega_\xi$ , one can interpret the first two principal components as coordinates. That is, we represent the covariation in our random coefficients on brand intercepts in two dimensions. Based on these coordinates, one can compute interbrand distances. Then we computed the congruence measure across maps—the correlation in interbrand distances obtained from the three specifications. Our results indicate that the implications for market structure are very similar across the models with a congruence measure between the Direct MLE and the two-step IV/MLE model of 0.96. Again, this finding implies that if the decisionmaker's focus is on understanding market structure then the standard random coefficients model will be well suited for the purpose.<sup>17</sup>

### 5.5. Targeting

We now use our results from the previous section to study household-specific tastes and the potential for targeting. Since the work of Rossi et al. (1996), marketers have developed practical means for using the distribution of consumer tastes to implement targeted marketing policies such as first-degree price discrimination. Our findings in the previous section

suggest that failure to account for unmeasured product attributes could have an adverse effect on marketing policy. In particular, the pricing manager might overestimate both the level of consumer willingness to pay and the degree of heterogeneity. In the current analysis, we consider the pricing decisions of a category manager.

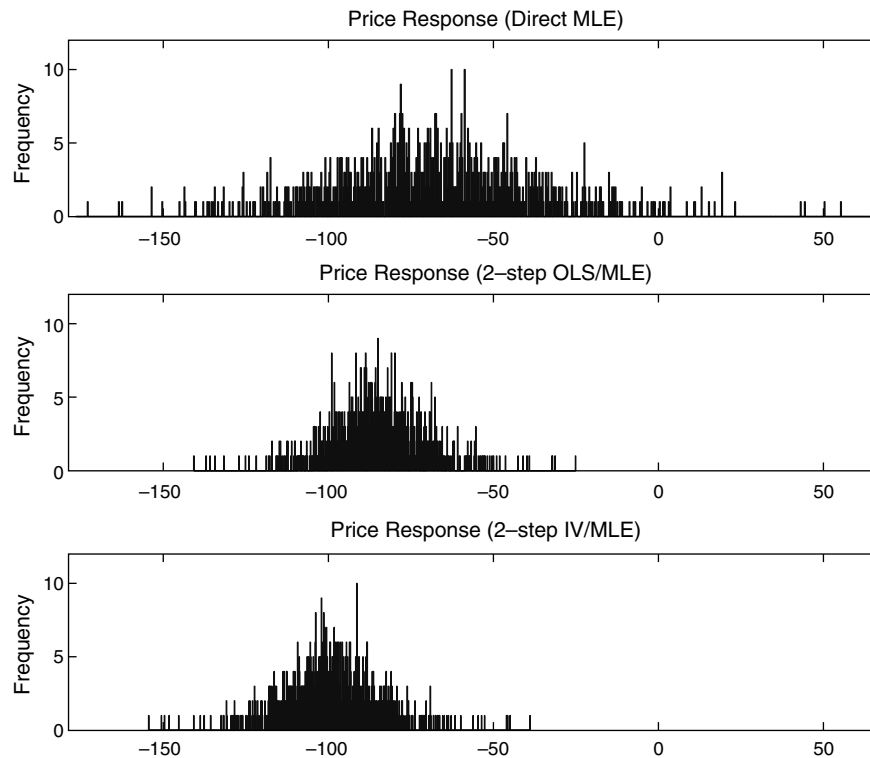
In Figure 1, we report the conditional distribution of consumer taste parameter based on each household's entire purchase history. For a given model, we draw 500 times from the asymptotic distribution of population tastes and then report the median for each household. We are now able to visualize the extent to which estimated heterogeneity diminishes with the addition of weekly fixed effects. The wide range of price-response parameters results in several households having very small in magnitude, often positive, price sensitivities under the simple heterogeneous model (Direct MLE). Both the potential endogeneity of prices and overstated heterogeneity may contribute to this problem. The former could bias the center of distribution toward the origin; the latter could widen the distribution so that the tails are widened on the positive portion of the support. Including fixed effects reduces the variance in tastes, tightening the tails of the distribution. Including instruments shifts the center of the distribution away from the origin. Hence, addressing the UBCs appears to resolve the practical concern insofar as positive price sensitivity parameters will prevent a manager from designing customized prices.

Also of practical concern is the role of the precision of our taste parameters once we include controls for UBCs and price endogeneity. In the results section, above in §5.2 we find that Direct MLE provide more precise, but seemingly biased estimates. We now check how precision in the population parameters influences our household conditional mean tastes. Although not reported in a table, using the same 500 draws from the asymptotic distribution for each model, we find the standard error of the household-specific price parameters rise from about 2.5, under Direct MLE, to about 8.5, under the two-step models. Despite the decrease in precision, we still find significant differences in these estimates in moving from Direct MLE to two-step IV/MLE. In particular, for over 70% of the households, the confidence intervals for the Direct MLE and two-step IV/MLE do not overlap. Thus, we obtain significantly more-price sensitive household estimates once we control for UBCs and price endogeneity.<sup>18</sup>

<sup>17</sup> Note that this finding is purely exploratory. Because we document it only in a single category, it is unclear whether a similar consistency would occur in other categories.

<sup>18</sup> Note that we only focus on the error around the household estimates due to error in the population taste parameters. We do not consider the error due to the variation in the household choice histories. This latter source of error is comparable across the three models because they all condition on the same histories.

Figure 1 Posterior Distribution of the Price Parameter



Note. For each model, we report each household's medians price parameters from a bootstrap with 500 replications from the population distribution of tastes.

As discussed previously, we do not simply plug the conditional taste parameters into the logit choice equation to compute targeted prices. Instead, we use the conditional purchase probability directly. Note that even with a negative expected price response parameter a household may still have a positive expected own-price elasticity, depending on the price level. This is an artifact of the primitive assumption that population tastes are distributed normally. Under the normal distribution, there is always some probability mass over the positive region of the support. As one integrates the heterogeneity out of the purchase probability equation via Monte Carlo simulation, one may obtain positive-valued draws. For households located closer to the origin of the price-response distribution, the own-price elasticity could be relatively high for positive draws and low for negative draws. As a result, the mean elasticity across draws may have a positive value. However, it will be impossible to compute targeted prices for households with positive price elasticities—the optimum is to set prices to infinity.<sup>19</sup>

<sup>19</sup> In practice, whenever a household has any positive-valued draws for the price parameter, the targeted pricing problem becomes indeterminate. A positive draw suggests there is a small probability that the consumer has positive price elasticity. As a result, the manager sets prices very high, moving the consumer onto the positive elasticity region of the expected demand.

In Table 8, we report the frequency of positive conditional expected price elasticity (e.g., the targeting problem cannot be solved). Nearly 30% of the households have indeterminate targeted prices under the Direct MLE model. After including weekly fixed effects, we still find 1% of the households with indeterminate prices. Only after instrumenting do we offset the problem entirely. In the case of Direct MLE, another solution to the positive price-parameter problem might be to make an ad hoc distributional assumption that restricts the entire probability mass to be negative, such as the log-normal or the truncated normal. The restricted distribution will provide a poorer fit of the data, because the unrestricted model predicts positive price response. Moreover, the restriction is somewhat arbitrary. In the current context, controlling for price endogeneity with instrumental variables solves the problem without the need for additional assumptions.

We also report the expected change in profits from switching from a model with optimal uniform pricing to one with targeted pricing. We report the mean profit improvement by household and overall. In both cases, the profit improvements fall considerably once we control for the presence of UBCs and the potential endogeneity of prices. In this example, we compare uniform to targeted pricing to illustrate the manner in which parameter biases due to omitted

**Table 8** Diagnostics of Targeting in Margarine Category

	No solution to targeted pricing (% households)	Change in profits from targeting (%)	
		Mean (per household)	Total (across households)
Direct MLE	0.27	0.14	0.19
Two-step OLS/MLE	0.01	0.03	0.04
Two-step IV/MLE	0.00	0.03	0.03

UBCs manifest themselves in managerial applications of the choice model. In practice, the uniform pricing model may also be inappropriate. In this case, one could compare profit improvements from switching between the observed pricing policy and some recommended policy, such as uniform or targeted pricing.

In Table 9, we also report the mean and standard deviation of the targeted prices for the nonindeterminate cases.<sup>20</sup> On average, the targeted prices appear similar across the three models; although the two-step IV/MLE produces lower prices than the two-step OLS/MLE model. As expected, the standard deviation in targeted prices is considerably lower once we include controls for UBCs. This result reflects the change in the extent of heterogeneity. Evidently, the main explanation for the lower profits under the two-step models appears to be related to heterogeneity and, to a lesser extent, to endogeneity.

## 6. Conclusions

We propose an empirical approach to accounting for the presence of unmeasured brand characteristics that could potentially be correlated with prices in a multinomial choice model. Our method allows for heterogeneity in preferences and on responses to marketing activities. Unlike previous approaches, we do not make strong assumptions regarding the nature of pricing behavior in the market or distributional assumptions regarding the UBCs. Our simulation results suggest that our method’s agnosticism regarding pricing behavior and the distribution of UBCs yields more robust results as compared to previous approaches.

We find strong evidence that ignoring the unobserved (to the econometrician) weekly brand-specific characteristics that influence consumer brand choice behavior may lead to higher estimated taste dispersion in a random coefficients logit model. We also find evidence consistent with the hypothesis that unmeasured brand characteristics affecting consumer choices may, in turn, affect retail pricing, inducing

<sup>20</sup> In the profit computations, we use the optimal uniform shelf price when the targeted price is indeterminate.

**Table 9** Mean and Standard Deviations of Targeted Prices for Margarine (Dollars per oz)

	BB		IC		PA		SH	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Het	0.70	0.96	1.35	0.97	0.85	0.96	0.89	0.87
Het, FE	0.75	0.08	1.28	0.08	0.86	0.09	0.84	0.06
Het, FE, IV	0.72	0.05	1.26	0.05	0.82	0.06	0.82	0.04

*Note.* BB, Blue Bonnet; IC, I Can’t Believe It’s Not Butter; PA, Parkay; SH, Shedd’s; Het, heterogeneous; FE, fixed effects.

correlation between prices and latent attributes. After instrumenting prices, we find a substantial increase in the predicted price-sensitivity of demand. Furthermore, we find evidence from the margarine category that failure to account for either endogeneity in prices or unobserved heterogeneity in consumer tastes will bias price sensitivity downward. We also find that loyalty remains significant after controlling both for heterogeneity and endogeneity. Comparable results are also available for the refrigerated orange juice category, on request. The results suggest that one must be careful to incorporate both these features into the estimation of household choice models. The combination of potentially overstated heterogeneity in tastes and biased mean responses implies that the posterior distribution of household level parameters is affected.

We discuss situations in which accounting for unobserved brand characteristics may be more or less important. For example, if the focus of the research is to identify the relative price elasticities across brands, then accounting for UBCs may be less important. Furthermore, if one is interested in characterizing market structure within a particular product category, ignoring the effects of UBCs may not be significant. However, for targeting, and in situations where the magnitudes of the elasticities are of interest, ignoring this feature of the model could lead to overconfidence in the ability to profit from one-to-one marketing policies.

In the current analysis, we have considered endogeneity only of prices. We hope that the simplicity of our approach encourages future marketing research to try to control endogeneity along other marketing instruments. Bajari and Benkard (2004) propose a nonparametric hedonic approach, in the spirit of Rosen (1974), that controls for unobserved product characteristics. This new line of research could be particularly interesting for marketers seeking to estimate more-flexible choice models, while being careful about the endogeneity of strategic marketing variables. More broadly, our evidence consistent with price endogeneity suggests the importance of pursuing more-active marketing research on the supply side



to try to learn more about the policies and rules of thumb used by managers to set marketing variables.

Finally, marketers seeking to estimate individual-level parameters with comparable household purchase data have typically used hierarchical Bayesian (HB) methods. The HB approach is generally believed to be better suited for recovering such estimates, especially given the small number of purchases typically available per household (Allenby and Rossi 1999). In general, there is no reason one could not conduct our analysis using such Bayesian techniques. A possible complication could be the selection of prior distributions for the model parameters. In particular, it would be challenging to construct sensible priors for the weekly brand fixed-effect parameters. Because these parameters include the weekly mean utilities for each brand, they most likely exhibit dependence across brands and weeks. Moreover, one would need to add more structure on the shocks in the second stage regression. Because prices are expected to be the outcome of some unspecified strategic decision process, imposing a parametric distribution could be restrictive. The advantage of the classical approach is that we do not need to make any assumptions about the joint distribution of the fixed effects. Also, for the case of the random effects logit, several studies have found HB and simulated maximum likelihood provide very similar individual parameter estimates.<sup>21</sup>

An online appendix to this paper is available at <http://mansci.pubs.informs.org/ecompanion.html>.

### Acknowledgments

The authors thank the editor, the associate editor, and two anonymous reviewers for their comments and suggestions. They have also benefited from the comments of Michaela Draganska, Jaehwan Kim, Puneet Manchanda, Alan Montgomery, Amil Petrin, Peter Rossi, Vishal Singh, Kannan Srinivasan, Miguel Villas-Boas, and Naufel Vilcasim, as well as seminar participants at Boston University, Carnegie Mellon Graduate School of —, the University of Chicago Graduate School of Business, the University of Colorado at Boulder, the Columbia Graduate School of Business, Dartmouth, Insead, Wharton, the 2002 Marketing Camp at Leuven, and the 2002 INFORMS conference on pricing at Cornell. The first two authors are grateful to the Kilts Center for Marketing for research support. P. Chintagunta thanks the Harvard Business School for its hospitality during the Fall 2002 semester.

### References

Allenby, G. M., P. E. Rossi. 1999. Marketing models of consumer heterogeneity. *J. Econometrics* **89** 57–78.

<sup>21</sup> For instance, Huber and Train (2001) find correlations of 0.98 between household parameters obtained from HB versus simulated MLE.

- Bajari, P., C. L. Benkard. 2004. Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. Working paper, Duke University, Durham, NC.
- Berry, S. 1994. Estimating discrete-choice models of product differentiation. *RAND J. Econom.* **25** 242–262.
- Berry, S., J. Levinsohn, A. Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* **63** 841–890.
- Besanko, D., S. Gupta, D. Jain. 1998. Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Sci.* **44**(November) 1533–1547.
- Blundell, R., J. Powell. 2001. Endogeneity in semiparametric binary response models. Working paper, University College London, London, U.K.
- Chiang, J. 1991. The simultaneous approach to the whether, what, and how much to buy questions. *Marketing Sci.* **10** 297–315.
- Draganska, M., D. C. Jain. 2002. Structural models of competitive market behavior. *Management Sci.* **50** 605–616.
- Dubé, J. P. 2004. Multiple discreteness and product differentiation: Demand for carbonated soft drinks. *Marketing Sci.* **23**(1) 66–81.
- Elrod, T., M. P. Keane. 1995. A factor-analytic probit model for representing the market structure in panel data. *J. Marketing Res.* **32** 1–16.
- Erdem, T., S. Imai, M. P. Keane. 2003. A model of consumer brand and quantity choice dynamics under price uncertainty. *Quant. Marketing Econom.* **1**(1) 1–65.
- Goolsbee, A., A. Petrin. 2004. The consumer gains from direct broadcast satellites and the competition with cable TV. *Econometrica* **72** 351–381.
- Guadagni, P. M., J. D. C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Sci.* **2** 203–238.
- Hajivassiliou, V., D. McFadden. 1993. The method of simulated scores for the estimation of LDV models with an application to external debt crises. Unpublished, University of California at Berkeley, Berkeley, CA.
- Huber, J., K. Train. 2001. On the similarity of classical and Bayesian estimates of individual mean partworths. *Marketing Lett.* **12** 259–269.
- Keane, M. 1993. Simulation estimation for panel data models with limited dependent variables. C. R. Rao, G. S. Maddala, D. Vinod, eds. *The Handbook of Statistics 11 Econometrics*. Elsevier Science Publishers, Amsterdam, The Netherlands, 545–572.
- Keane, M. P. 1997. Modeling heterogeneity and state dependence in consumer choice behavior. *J. Bus. Econom. Statist.* **15** 310–327.
- Kim, J., G. M. Allenby, P. E. Rossi. 2002. Modelling consumer demand for variety. *Marketing Sci.* **21** 229–250.
- Kuksov, D., J. M. Villas-Boas. 2003. Endogeneity and individual consumer choice. Working paper, Washington University, St. Louis, MO.
- McFadden, D. 1981. Econometric models of probabilistic choice. C. Manski, D. McFadden, eds. *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge, MA.
- McFadden, D., K. Train. 2000. Mixed MNL models for discrete responses. *J. Appl. Econometrics* **15** 447–470.
- Mehta, N., S. Rajiv, K. Srinivasan. 2003. Price uncertainty and consumer search: A structural model of consideration set formation. *Marketing Sci.* **22** 58–84.
- Nevo, A. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica* **69**(2) 307–342.
- Petrin, A., K. Train. 2002. Omitted product attributes in discrete choice models. Working paper, University of Chicago, Chicago, IL.

- Revelt, D., K. Train. 1999. Customer-specific taste parameters and mixed logit. Working paper, University of California at Berkeley, Berkeley, CA.
- Rosen, S. 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. *J. Political Econom.* **82** 34–55.
- Rossi, P. E., R. E. McCulloch, G. M. Allenby. 1996. The value of purchase history data in target marketing. *Marketing Sci.* **15** 321–340.
- Sudhir, K. 2001. Structural analysis of competitive pricing in the presence of a strategic retailer. *Marketing Sci.* **20** 244.
- Van Nierop, E., R. Paap, B. Bronnenberg, P. H. Franses, M. Wedel. 2001. Modeling unobserved consideration sets for household panel data. Working Paper 356, Marketing Studies Center, John E. Anderson Graduate School of Management, University of California at Los Angeles, Los Angeles, CA.
- Villas-Boas, M., R. Winer. 1999. Endogeneity in brand choice models. *Management Sci.* **45**(10) 1324–1338.
- Villas-Boas, M., Y. Zhao. 2001. The ketchup marketplace: Retailer, manufacturers, and individual consumers. Working paper, University of California at Berkeley, Berkeley, CA.
- Yang, S., Y. Chen, G. Allenby. 2003. Bayesian analysis of simultaneous demand and supply. *Quant. Marketing Econom.* **1**(3) 244–264.