Research Note
Endogeneity and Heterogeneity in a Probit Demand Model: Estimation Using Aggregate Data

Pradeep K. Chintagunta
Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, Illinois 60637
pradeep.chintagunta@gsb.uchicago.edu

Abstract

Two issues that have become increasingly important while estimating the parameters of aggregate demand functions to study firm behavior are the *endogeneity* of marketing activities (typically, price) and *heterogeneity* across consumers in the market under consideration. Ignoring these issues in the estimation of the demand function parameters can lead to biased and inconsistent estimates for the effects of marketing activities. Endogeneity and heterogeneity have achieved prominence in large measure because of the increasing popularity of logit models to characterize demand functions using *aggregate* data. The logit model accounts for purchase incidence and brand choice by including a ''no-purchase'' alternative in the consumer's choice set. This allows for category sales to change as a function of the marketing activities of brands in the category.

There are three issues with using the logit model with the no-purchase option to characterize demand when studying competitive interactions among firms. (1) The marketing literature dealing with brand choice behavior at the consumer level has found that the IIA restriction is not appropriate, as each brand in the choice set is more similar to some brands than it is to others. (2) Studies have found that the purchase incidence decision is distinct from the brand choice decision. Hence, it may not be appropriate to model the no-purchase decision as just another alternative in the choice set with the IIA restriction holding across all brands and the no-purchase option. (3) Even if the distinction between the purchase incidence and brand choice decisions is accounted for via, for example, a nested logit specification, accounting for the purchase incidence decision with aggregate data requires assumptions for computing the share of the no-purchase alternative which is otherwise unobserved.

In this paper, we propose a probit model as an alternative to the logit model to specify the aggregate demand functions of firms competing in oligopoly markets. The probit model avoids the IIA property that affects the logit model at the individual consumer level. Furthermore, the probit model can naturally account for the distinction between the purchase incidence and brand choice decisions due to the general covariance structure assumed for the utilities of the alternatives. We demonstrate how the parameters of the proposed model can be estimated using aggregate time series data from a product market. In the estimation, we account for the endogeneity of marketing variables as well as for heterogeneity across consumers.

Our results indicate that both endogeneity as well as heterogeneity need to be accounted for even after allowing for a non-IIA specification at the individual consumer level. Specific to our data, we also find that ignoring endogeneity has a bigger impact on the estimated price elasticities than ignoring the effects of heterogeneity. A comparison of the elasticities obtained from the probit model with those from the corresponding logit specification indicates that the *range* of elasticities obtained from the probit model across brands is larger than that obtained from the logit. The results have implications for issues such as firm-level pricing.

In addition to specifying a probit model and providing comparisons with the logit model, the paper also addresses the third issue raised above. We propose a simple alternative to the purchase incidence/brand choice specification by decomposing the demand for a brand into a category demand equation and a conditional brand choice share equation. We provide a comparison of results from this specification to those from the specification that includes the no-purchase alternative and find that estimated elasticities are sensitive to the specification used. We also estimate the demand function parameters using a traditional specification such as the double-logarithmic model. Here, we find that the estimated elasticities could be signed in such a manner as to be not useful for firm-level pricing decisions.

One of the key limitations of the proposed model is that while it accounts for the purchase incidence and brand choice decisions of households, it does not account for differences across consumers in their purchase quantities. The model and analysis are best suited for product categories in which consumers typically make single-unit purchases. Another limitation is more practical in nature. While recent advances have been made in computing probit probabilities, it could nevertheless be a challenge to do so when the number of alternatives is large.
(*Heterogeneity; Endogeneity; Probit Model; Logit Model)*

MARKETING SCIENCE © 2001 INFORMS
Vol. 20, No. 4, Fall 2001, pp. 442–456

0732-2399/01/2004/0442/$05.00
1526-548X electronic ISSN

Introduction

The recent literature in marketing and in economics has seen an explosion of studies dealing with the analysis of firm-level behavior. The principal motivation behind these studies is to measure market power of firms and to understand interfirm competitive behavior. For example, Kadiyali (1996) studies the competitive interactions between Kodak and Fuji in the U.S. market to investigate whether or not the rivalry between these two firms is as intense as indicated by the popular press. Sudhir (2001) looks at the competitive interactions among firms within various segments of the automobile industry to determine whether these interactions vary significantly across product segments. Nevo (2001) investigates the extent of pricing rivalry in the ready-to-eat breakfast cereal market to determine whether observed prices reflect market power associated with product differentiation or collusion by firms in the industry.

The fundamental building block for the analysis of firm behavior is the demand function for each of the players in the marketplace. The demand functions relate the sales of the brands to their prices, promotions, and other marketing variables. Two issues that have become increasingly important while estimating the parameters of such aggregate demand functions to study firm behavior are the *endogeneity* of marketing activities (typically, price) and the *heterogeneity* across consumers in the market under consideration. The endogeneity problem arises when there are variables for which data are not available (such as shelf space allocation, shelf location, store coupons, etc.) that could influence a brand's sales in a given week and if these variables are correlated with the included marketing variables such as price (lowering the price of brand in a given week may be accompanied by giving it more shelf facings). These other marketing activities are part of the error term in the estimation and the correlation between the price variable and the error term results in the endogeneity problem. Not accounting for this correlation will give incorrect estimates for the effects of the included marketing variables. The issue with heterogeneity is the same as it is with household data. If the observed data at the store or market level are the aggregation of consumers with different brand preferences and sensitivities to marketing instruments, then ignoring this heterogeneity in the estimation of the demand function parameters can lead to biased and inconsistent estimates for the marketing activities.

The issues of endogeneity and heterogeneity have achieved prominence in large measure because of the increasing popularity of discrete-choice models to specify the demand functions to study firm behavior. The aforementioned studies by Sudhir and Nevo, along with others by Berry et al. (1995), have used discrete-choice-based demand functions. The main advantages of discrete-choice models are: (i) They are derived from utility maximizing behavior of consumers in the marketplace. (ii) They require estimation of fewer numbers of parameters, as compared to linear (and log–log and semilog) demand functions (instead of estimating 100 price parameters in a market with 10 brands, usually a single price parameter is estimated). (iii) They seldom result in parameter estimates with incorrect signs for own and cross effects, as is the case with linear demand systems and their variants.

The most widely used specification of discrete-choice demand function in such studies thus far has been the logit model. To allow the total demand for the category to vary over time, the model treats the no-purchase option or the "outside good" as an additional alternative available in the choice set. The specification embodies all the advantages of discrete-choice models noted above. Additionally, it is also easy to estimate. All these advantages appear to justify the model's widespread use in the literature. Hence, researchers have used the logit demand model and have accounted for the issues of endogeneity and heterogeneity while estimating the parameters of these models with aggregate store (Besanko et al. 1998), chain, or market (Sudhir 2001) data. Accounting for heterogeneity in the logit model also alleviates the problem of restrictive cross-elasticities that are obtained from this model because of the "independence of irrelevant alternatives" (IIA) property at the individual consumer level (see the discussion in Nevo 2001).

While the logit model with an outside good (and accounting for endogeneity and heterogeneity in the estimation with *aggregate* data) has seen widespread application in the marketing and economics literature, little research has been devoted to analyzing the sensitivity of the results obtained to some of the restrictions associated with this specification. Our goal in this paper is to investigate two of these restrictions. The first is the IIA restriction of the logit model at the individual consumer level. Models such as the probit (Currim 1982) do not suffer from the IIA problem at the individual level. As noted above, *aggregate* elasticities from the logit model that accounts for heterogeneity are indeed not subject to the IIA restriction. Nevertheless, researchers who have estimated logit and probit models with household data have found the aggregate elasticities from these two specifications to be different even after accounting for the effects of heterogeneity (see Chintagunta and Honore 1996). So the question that arises is: Are the aggregate elasticities obtained from the logit and probit models similar when using aggregate data in the estimation and after accounting for the effects of both heterogeneity and endogeneity? An answer to this question is important, as the elasticities directly influence the measure of market power.

The second issue we investigate is the modeling of the ''no-purchase'' option as an additional alternative in the logit model. Inclusion of the outside good in the estimation requires the shares of each of the alternatives—including that for the outside good or ''no-purchase'' alternative—to be known. With consumer level data, where one observes whether or not a household purchases the product category, computing the shares is straightforward (Chintagunta 1993). With aggregate data, we only observe the sales or shares of the brands but do not observe the aggregate fraction of consumers not buying a product category in a given week. Hence, we need to assume the total potential size of the market in each week to compute the share of the outside good (Nevo 2001 assumes that everyone living in the market area consumes the equivalent of one helping of cereal each day). In this paper, we present a simple alternative to the approach used in previous studies. Rather than attempt to model the purchase incidence and brand choice decisions simultaneously, we employ the approach proposed by Kim et al. (1995). Category sales are modeled as a regression of the total demand across brands on category level marketing activities. Brand shares are obtained as an aggregation of consumers' conditional (on category purchase) brand choice probabilities as in previous studies such as those by Berry et al. (1995), Nevo (2001), Sudhir (2001), etc. The advantage of this methodology is that we only use information on the brands that is directly obtained from the marketplace, i.e., sales, prices, and other marketing activities of brands. A disadvantage is that the model can no longer be given a fully ''structural'' interpretation, as the category regression is a reduced-form approach to modeling a piece of the consumer's decision problem (the purchase incidence decision). We provide a comparison of results obtained from the two methods for our data to investigate the sensitivity of the elasticity estimates to the definition of the outside good.

The remainder of this paper is organized as follows. In the next section, we describe the estimation of the aggregate probit model that accounts for heterogeneity as well as price endogeneity. We then provide the results from the empirical analysis using market data on shampoo purchases. A comparison with the logit specification and an investigation of the sensitivity to alternative assumptions on the no-purchase option are provided. The final section concludes with some directions for future research using the methodology.

Model Formulation and Estimation Strategy

We begin with the basic probit model at the household (we use consumer and household interchangeably here) level. We then describe the category level regression model. Our description of the probit model assumes the presence of K ''brands.'' This can be interpreted as $K - 1$ brands and one no-purchase option (in the case where the outside alternative is part of the choice process) or as K brands with a separate category sales regression. In the former case, the Kth brand will not have any marketing variables associated with it. Specifically, the indirect utility of consumer i for brand j in week t is given by the following expression.

$$V_{ijt} = \alpha_{ij} + \beta_i \ln(p_{jt}) + \gamma d_{jt} + \mu_{jt} + \epsilon_{ijt},$$
$$V_{ijt} = Y_{ijt} + \epsilon_{ijt}, \tag{1}$$

where the category consists of K brands $j = 1, 2, 3, \ldots, K$. V_{ijt} is the indirect utility of consumer i for brand j in week t. ϵ_{ijt} is a K-variate normal random error term with mean 0 and covariance matrix Ω. Y_{ijt} includes all terms in the indirect utility function excluding ϵ_{ijt}. α_{ij} is consumer i's intrinsic preference for brand j. β_i is the price sensitivity parameter for consumer i. p_{jt} is the price of brand j in week t. γ is the deal parameter and d_{jt} is the deal variable for brand j in week t. μ_{jt} is the unobservable attribute for brand j in week t. The unobservable attribute μ_{jt} captures the effects of variables other than prices and deals that are not included in the model and that could drive the probability of choosing brand j. These are in-store variables that could vary over time and are correlated with the retail price. This results in the endogeneity problem discussed in the introduction. The probability of brand j being chosen is given by:

$$\begin{aligned} P_{ijt} &= \Pr(V_{ikt} - V_{ijt} \le 0, \forall k = 1, 2, \ldots, K, k \ne j) \\ &= \Pr\{\epsilon_{ikt} - \epsilon_{ijt} \le \alpha_{ij} - \alpha_{ik} + \beta_i[\ln(p_{jt}) - \ln(p_{kt})] \\ &\qquad + \gamma(d_{jt} - d_{kt}) + \mu_{jt} - \mu_{kt}, \\ &\qquad k = 1, 2, \ldots, K, k \ne j\} \\ &= \Pr(\eta_{j,ikt} \le Z_{j,ikt}, k = 1, 2, \ldots, K, k \ne j) \\ &= \Phi(\mathbf{Z}_{ijt}, \Omega_{j,K-1}) \end{aligned} \tag{2}$$

where η_{ikt} has a $(K - 1)$-variate normal distribution with mean zero and covariance matrix $\Omega_{j,K-1}$, $\Phi(.,.)$ refers to the CDF of a $K - 1$ variate normal distribution, and $\mathbf{Z}_{ijt}$ denotes the matrix $Y_{ijt} - Y_{ikt}$ $\forall k = 1, 2, \ldots, K, k \ne j$ with each element denoted by $Z_{j,ikt}$ $\forall k = 1, 2, \ldots, K, k \ne j$. Note that, unlike the logit model, the probability in Equation (2) does not have a closed form expression and represents a $K - 1$ dimensional integral. However, there are several approaches to computing the integral to a high degree of accuracy. See Hajivassiliou et al. (1996) for a comparison of these methods.

In the above formulation, households are assumed to differ in their preferences as well as in their price sensitivities. To account for such heterogeneity, researchers have proposed several approaches. The two most commonly used specifications are the parametric random effects logit model (Allenby and Rossi 1999) or the semiparametric random effects logit model (see for example, Jain et al. 1994). Here, we focus on the parametric model. For the latent class approach using aggregate data, see Berry et al. (1998). Heterogeneity in intrinsic preferences (α_{ij}) and price sensitivities (β_i) are accounted for as follows:

$$\alpha_{ij} = \alpha_j + \epsilon_{ij} \quad \text{and} \quad \beta_i = \beta + \epsilon_{i\beta},$$

where

$$\epsilon_{ij} \sim N(0, \sigma_j^2) \quad \text{and} \quad \epsilon_{i\beta} \sim N(0, \sigma_\beta^2). \tag{3}$$

α_j is the mean intrinsic preference level for brand j across households, and β is the mean value of the price sensitivity parameter. The term σ_j^2 represents the variance in the intrinsic preference for brand j across consumers. With household level data, one can allow these preferences to be correlated across brands. However, with aggregate data, we do not have information to distinguish between these correlations and those due to the random component of indirect utilities, ϵ_{ijt}. Hence, while we assume utilities themselves to be correlated across brands, we restrict the preferences to be uncorrelated across brands. Three points are noteworthy at this juncture.

(i) If the brands or products included in the analysis can be represented by their constituent attributes, then allowing for heterogeneity along each attribute as in Equation (3), will allow for brand preferences to be correlated without the problem of identification noted above.

Specifically, let $\alpha_{ij} = \sum_{w=1}^{W} \alpha_{iw} I_{jw}$ where I is an indicator that takes the value 1 if brand or product j includes attribute w or zero otherwise, and α_{iw} is the preference value that consumer i has for attribute w. Then, Equation (3) can be written as $\alpha_{ij} = \Sigma_{w=1}^{W} \alpha_w I_{jw} + \Sigma_{w=1}^{W} \epsilon_{iw} I_{jw}$. Even if the ϵ_{iw} terms are independent across the w's, the preferences for the two brands j and k will be correlated if they share a subset of attributes. By imposing a structure on the nature of preference correlation, we can overcome the identification problem noted previously.

(ii) If we had access to data from multiple markets or multiple stores in a given market, we can exploit in addition, the variation in demographic characteristics across the different units (markets or stores) by making α_{ij} and β_i functions of these variables. In this way, we can allow for systematic differences in preferences as well as sensitivities to marketing activities across different demographic units.

(iii) If the probabilities in (2) are based on the logit model, we can allow for a general pattern of correlation across preferences and price sensitivities in (3), as the utilities themselves are constrained to be uncorrelated in this case.

When one has access to household data (in the absence of the unobserved attribute term μ_{jt}), we can write out the likelihood of a string of purchases over time for each household. This likelihood would then be integrated over the distribution of heterogeneity. The sample likelihood, which is the product of the unconditional household likelihoods, would then be maximized to arrive at a set of parameters. Villas-Boas and Winer (1999) have recently addressed the issue of accounting for μ_{jt} and then estimating the parameters of the logit demand model using household data. In dealing with aggregate data, estimation is complicated by two issues: (a) Data are observed only at the aggregate level. In other words, what we observe are S_{jt}—shares of brand j in week t. (b) p_{jt} and μ_{jt} could potentially be correlated.

The principle underlying the estimation is simple: Obtain estimates that equate the observed shares S_{jt} to the shares predicted by the model, s_{jt}. The implementation of this strategy tends to be more complicated because of the correlation mentioned above. The approach used closely parallels that of Berry et al. (1995) and Nevo (2001), with one important difference (that we discuss subsequently).

Step 1. Decompose Y_{ijt} as $Y_{ijt} = (\alpha_j + \beta \ln(p_{jt}) + \gamma d_{jt} + \mu_{jt}) + [\epsilon_{ij} + \epsilon_{i\beta} p_{ijt}] = L_{jt} + [\epsilon_{ij} + \epsilon_{i\beta} p_{ijt}]$. Note that L_{jt} is household invariant, whereas the second term depends on i. Intuitively, the estimation involves two ''nested'' loops. In the ''outer'' loop, the parameters corresponding to the household heterogeneity distribution as well as those in Ω_{K-1} (Equation 2) are computed, whereas the ''inner'' loop involves computing the unknown parameters embedded in L_{jt}. It is important to distinguish between the two loops, because while L_{jt} is linear in the unknown parameters, α_j, β, and γ, the term in the square brackets is nonlinear in the embedded parameters (Ω_{K-1}, σ_j^2 and σ_β^2).

Step 2. Make R draws for the terms ϵ_{ij} and $\epsilon_{i\beta}$. This requires initial guesses for the unknown (nonlinear) parameters σ_j^2 and σ_β^2. Hence, given these initial values, the term in the square brackets in the above equation is ''known.'' Additionally, in this step, we also need to pick starting values for the parameters in Ω_{K-1} (to ensure that the matrix is positive definite, we choose initial values for the Cholesky decomposition of this matrix).

Step 3. Make initial guesses for the L_{jt} terms. Note that if there are 3 brands and 100 time periods, this involves 300 L_{jt} ''parameters'' in the case of the outside good model and 200 for the category sales/brand choice model. Now, given L_{jt}, $[\epsilon_{ij} + \epsilon_{i\beta} p_{ijt}]$, and the starting values for Ω_{K-1}, we can compute the probit probability P_{rjt} for each of the R draws. The predicted share from the model (s_{jt}) is the average probability across the R draws.

Step 4. The ''inner loop'' computation takes place. In other words, keeping the nonlinear parameters fixed at the initial guesses, we iterate over the (200 or 300) values of L_{jt} to minimize the distance between the predicted share (s_{jt}) and the actual share (S_{jt}). Given the nonlinearity of the probit probability, the logarithmic transformation to linearity (see Berry et al. 1995) that works for the logit model no longer applies, and we need to use standard nonlinear optimization methods to carry out the minimization. This step can be summarized as follows (note L_t is the vector $\{L_{1t}, L_{2t}, \ldots, L_{Kt}\}$).

$$\min_{L_t} (S_{jt} - s_{jt}) = \min_{L_t} [S_{jt} - \Phi(L_t | \mathbf{Z}_{jt}, \sigma_j, \sigma_\beta, \Omega_{j,K-1})],$$

$$\min_{L_t} \Bigg[S_{jt} - \int_{\eta_{j,Kt} \le Z_{j,Kt}} \cdots \int_{\eta_{j,(j+1)t} \le Z_{j,(j+1)t}} \int_{\eta_{j,(j-1)t} \le Z_{j,(j-1)t}} \cdots \times \int_{\eta_{j,1t} \le Z_{j,1t}} \phi(\eta_{j,1t}, \ldots, \eta_{j,(j-1)t}, \eta_{j,(j+1)t}, \ldots, \eta_{j,Kt})\, d\eta_{j,1t}, \ldots, \eta_{j,(j-1)t}, \eta_{j,(j+1)t}, \ldots, \eta_{j,Kt} \Bigg].$$

Step 5. Step 4 gives us the values L_{jt} for all j and t. Now, returning to the expression for L_{jt}, we note that $L_{jt} = \alpha_j + \beta \ln(p_{jt}) + \gamma d_{jt} + \mu_{jt}$. If $\mathrm{corr}(p_{jt}, \mu_{jt}) = 0$, then we can obtain α_j, β, and γ by simply regressing L_{jt} on intercepts, $\ln(p_{jt})$ and d_{jt}. However, given the possibility of correlation, instrumental variable methods are used instead. This completes the computation of the linear parameters, *conditional* on the *initial* choices of the nonlinear parameters.

Step 6. The error term μ_{jt} is computed as $L_{jt} - (\alpha_j + \beta \ln(p_{jt}) + \gamma d_{jt})$.

Step 7. The error term is then interacted with the instrument vector used in Step 5 to provide the GMM objective function. This objective function forms the basis of obtaining the nonlinear parameters, i.e., the outer loop.

Step 8. Minimizing the GMM objective function by iterating over the values of Ω_{K-1}, σ_j^2, and σ_β^2 provides estimates for the nonlinear parameters. The corresponding values of α_j, β, and γ computed in Step 5 will give us the values of the linear parameters. The standard errors of the estimates can then be computed. We turn next to the formulation of the category regression model in instances in which it is difficult to quantify the sales of the outside good.

The Category Sales Regression

In the case of the probit model specification with $K - 1$ brands and no outside good, prices of the various brands have no influence on the total size of the category. To overcome this problem, we propose coupling the probit brand choice model with a category sales model. Denote by Q_{jt} the sales of brand j in week t. Then the sales at the "category" (or subcategory level in our case) is nothing but the aggregation of sales across brands. The category sales in week t is given by $CQ_t = \sum_{j=1}^{K-1} Q_{jt}$. The category sales level will depend on the prices and promotions of the various brands in the category and also on factors such as seasonality. We compute category level price and promotion variables by share-weighting the prices and promotions of the individual brands (see Kim et al. 1995). Rather than use a weekly share weight however, we compute the average share of each brand over the period of the data and use these as share weights. Therefore, variation in the dependent variable is not being used to create our independent variables. We denote the share-weighted price and promotion variables as CP_t and CR_t. Now the category sales regression model is given as follows:

$$\ln(CQ_t) = \omega + \nu \ln(CP_t) + \rho CR_t + \sum_{s=1}^{3} \lambda_s I_{st} + e_t. \quad (4)$$

In the above equation, ω, ν, ρ, λ_s are parameters to be estimated. I_{st} is an indicator variable taking the value 1 if week t is in season s and zero otherwise. The random error term is e_t. Estimation of the parameters of the above equation requires recognizing two important points. The first is that the category price is likely to be endogenous, i.e., potentially correlated with the error term. Furthermore, e_t could be correlated with μ_{jt} from Equation (2). The first issue can be addressed by using instruments for category pric-

Table 1 Descriptive Statistics

	Brand A		Brand B		Brand C	
Variable	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Sales (units)	906,697	132,412	1,316,436	1,557,747	1,639,378	214,326
Share	0.235	0.024	0.341	0.029	0.424	0.036
Price (16-oz. bottle)	2.562	0.094	2.561	0.112	2.109	0.019
Promotion	1,790	6,892	622	2,525	8,822	18,396

es in addition to instruments for brand prices. The second issue can be addressed by augmenting the error vector μ_t used in the GMM with the error term e_t. The prediction of the aggregate sales of brand j in week t is obtained by multiplying the prediction from the category sales model with that obtained from the brand choice model. We have the following expression:

$$Q_{jt} = \left\{\exp\left[\omega + \nu \ln(CP_t) + \rho CR_t + \sum_{s=1}^{3} \lambda_s I_{st} + e_t\right]\right\} \times \left[\frac{1}{R}\sum_{r=1}^{R} \Phi(\mathbf{Z}_{rjt}, \Omega_{j,K-1})\right]. \quad (5)$$

Note that it is important to account for the error terms e_t and μ_{jt} when making predictions. This is similar in kind to the issue raised by Christen et al. (1997) in the context of log–log regression models.

In the case in which the no-purchase option or the outside good is treated as an additional alternative, the sales of brand j in week t are given by the following expression. M in the equation refers to the total consumption associated with that category in each week (assumed to be invariant over time). For example, as described above, in the case of Nevo (2001), this is the potential consumption of cereal by the population of interest.

$$Q_{jt} = M\left[\frac{1}{R}\sum_{r=1}^{R} \Phi(\mathbf{Z}_{rjt}, \Omega_{j,K})\right]. \quad (6)$$

Note from the above equation that the covariance matrix $\Omega_{j,K}$ is of dimension K, as opposed to $K - 1$, as there is one additional alternative—the outside good.

Estimation and Results

The data we use are for the shampoo product category. Because of the proprietary nature of the data, we are unable to reveal the actual identities of the brands. There are three brands in the specific subcategory chosen for the analysis and we refer to them as brands A, B, and C. We chose this subcategory because the managers at the firm releasing the data felt that these brands formed a distinct submarket in the category. The data are aggregated for the entire U.S. market. While it is important to consider issues of aggregation as described in Christen et al. (1997), market level data are routinely used for the investigation of competitive interactions. Weekly information over two years (104 weeks) is available for the three brands. Besides the sales levels of the brands, we also have their levels of prices and promotional activities over the 2-year period. In addition, we used seasonal dummies in the category sales regression. We assume that the only endogenous variable is price. The instruments we used are the following. From the Bureau of Labor Statistics, we obtained price indices for material (packaging as well as certain categories of chemicals used as ingredients in shampoos) and labor. We also used values of one period lag prices for all brands as instruments. Note that lagged prices can be problematic when there is serial correlation in the μ_{jt} term.

Descriptive statistics of the data are in Table 1. The share data are conditional on purchase. They indicate that brand C is the biggest brand in this particular subcategory of the shampoo category. However, the smallest brand, brand A has the highest coefficient of variation of the three brands. The average prices of brands A and B are very close in magnitude to each other, although there appears to be greater variation in the prices of brand B. Brand C, the largest share brand has the lowest price. As this category is heavily promoted through manufacturer coupons, we use information on the couponing variable to capture promotional effects on sales. The variable is operationalized as the total value of coupons dropped in each week. Table 1 indicates that brand C drops the most coupons, followed by brands A and B. The low price of brand C coupled with its heavy couponing appear to contribute to its large share in the marketplace. In the estimation, we used current and lagged values of the couponing variable. We found that the only significant variable was the 1-week lagged value of coupons dropped. Hence, this is the only variable included in the subsequent estimation and results.

Estimation

In the estimation, we performed extensive sensitivity analyses pertaining to the number of draws from the heterogeneity distribution required. Based on this, we settled on 100 draws ($R = 100$) as being reasonable, as increasing the draws beyond this number did not affect the parameter estimates significantly. For the outside good models, we did not use any marketing variables in the utility specification. If data are available, they can easily be incorporated into the analysis.

Results

The results are discussed as follows. First, we discuss the estimates obtained from the category regression/brand choice probit model. Next, we provide results from the probit model with an outside good included in the specification but without the category regression model. Finally, we discuss the results obtained from the comparison logit models. In Columns 2–5 of Table 2, we provide the results from the probit model with the category regression/brand choice formulation. Four different specifications were estimated. These are (i) without the unobserved attribute μ_{jt} that results in the endogeneity problem and without accounting for heterogeneity; (ii) accounting for the effects of endogeneity but not for the effects of heterogeneity; (iii) without the unobserved attribute μ_{jt} but accounting for heterogeneity in preferences as well as the price sensitivity parameter; and (iv) the most general case that accounts for both endogeneity as well as heterogeneity.

From Table 2, we see that brand A is the designated ''base'' brand with mean intrinsic preference level set to zero. The two specifications that account for endogeneity reveal positive mean intrinsic preferences for the two larger brands, B and C, as their estimates exceed zero. In the models that do not account for endogeneity, brand B has a lower mean intrinsic preference level than brand A. Note that the magnitudes of these and other estimates are not directly comparable because of differences in the estimated covariance matrices across the four specifications. Table 2 also reveals that the coefficients of the two marketing variables, price, and promotion have the right signs and are significant at the 5% level of significance across all the model specifications. In order to interpret the relative magnitudes of the price coefficients, we compute the corresponding elasticities that are presented later. The heterogeneity parameters in Table 2 indicate that there is some heterogeneity in the intrinsic brand preferences in the case of the most general model although the variances are not very large in magnitude. This implies that after one explicitly allows for non-IIA behavior at the individual consumer level via the probit specification, there appears to be little heterogeneity in intrinsic preferences. Later, we will contrast this finding with that obtained from the corresponding logit model specifications. The most heterogeneity we find is for the price sensitivity parameter obtained under the ''with endogeneity and with heterogeneity'' specification. In this case, the standard deviation of 0.242 is significantly different from zero. Note that there are three covariance parameters estimated as there are three brands. Hence, $\Omega_{j,K-1}$ is a 2×2 matrix with three unknown parameters in the Cholesky decomposition, of which only two parameters are uniquely identified. Hence, standard errors are not reported for the third parameter, as it is fixed.

Table 2 Parameter Estimates and Standard Errors for Shampoo Data*

	Category Regression/Brand Choice Model					Outside Good Model	
Variable	No Endogeneity No Heterogeneity	With Endogeneity No Heterogeneity	No Endogeneity With Heterogeneity	With Endogeneity With Heterogeneity	Variable	With Endogeneity No Heterogeneity	With Endogeneity With Heterogeneity
Brand B	−0.204	0.028	−0.150	0.034	Brand A	−0.010	0.017
	(0.024)	(0.031)	(0.026)	(0.032)		(0.004)	(0.005)
Brand C	0.114	0.007	0.062	0.034	Brand B	0.002	0.032
	(0.019)	(0.025)	(0.021)	(0.026)		(0.005)	(0.004)
Price	−0.570	−0.298	−0.404	−0.428	Brand C	−0.007	0.009
	(0.121)	(0.137)	(0.120)	(0.135)		(0.003)	(0.004)
Promotion**	0.007	0.004	0.005	0.004	Price	−0.142	−0.171
	(0.001)	(0.002)	(0.001)	(0.002)		(0.059)	(0.062)
Covariance Parameter 1	0.460***	0.083***	0.236***	0.054***	Promotion	0.003	0.003
						(0.001)	(0.001)
Covariance Parameter 2	1.383	0.287	0.960	0.279	Covariance Parameter 1***	0.196	0.179
	(0.079)	(0.111)	(0.113)	(0.128)			
Covariance Parameter 3	0.548	0.305	0.383	0.384	Covariance Parameter 2	−0.167	−0.167
	(0.082)	(0.129)	(0.095)	(0.134)		(0.079)	(0.081)
σ_β (Price)	—	—	0.049	0.242	Covariance Parameter 3	0.001	−0.0001
			(0.021)	(0.029)		(0.065)	(0.071)
σ_A (brand A)	—	—	0.008	0.031	Covariance Parameter 4	0.252	0.240
			(0.010)	(0.014)		(0.081)	(0.111)
σ_B (brand B)	—	—	0.016	0.005	Covariance Parameter 5	0.145	0.144
			(0.012)	(0.015)		(0.066)	(0.051)
σ_C (brand C)	—	—	0.007	0.048	Covariance Parameter 6	0.143	0.134
			(0.009)	(0.016)		(0.071)	(0.066)
					σ_β (Price)	—	0.058
							(0.021)
Intercept**	0.161	0.159	0.161	0.160	σ_A (brand A)	—	0.004
	(0.033)	(0.056)	(0.033)	(0.056)			(0.032)
Price	−0.011	−0.009	−0.011	−0.010	σ_B (brand B)	—	0.073
	(0.002)	(0.003)	(0.003)	(0.003)			(0.022)
Promotion**	0.0002	0.0002	0.002	0.0002	σ_C (brand C)	—	0.027
	(0.00005)	(0.0001)	(0.00006)	(0.0001)			(0.027)

*The estimates for the seasonality parameters are not reported, as they were not significant at the 5% level.
**Promotion variable was multiplied by 1e−4 and log(Category Sales, Table 3) was multiplied by 0.01 in the estimation.
***Fixed in the estimation. Note that only $(K \cdot (K - 1)/2) - 1$ parameters of the covariance matrix are identified.

Table 2 provides the parameter estimates obtained from the category regression model under each specification. A priori, we would expect the parameters from the two "no-endogeneity" specifications to resemble each other and those from the two "with-endogeneity" models to be similar as heterogeneity has no impact on the category sales regressions. Indeed the results reflect this, although results from all four specifications are quite similar to one another. One of the things we also find is that seasonality does not play a major role in this product category.

In Table 3, Columns 2–9, we present the elasticity estimates from the four specifications. For each specification, we present two sets of elasticities. The first column corresponds to the brand choice elasticities. The second column contains the total sales or de-

Table 3 Price Elasticities (Standard Errors) from the Various Models

	Probit Models								Logit Models			
	Category Regression/Brand Choice											
	No Endogeneity & No Heterogeneity		Only Endogeneity		Only Heterogeneity		Endogeneity & Heterogeneity		Outside Good		Endogeneity & Heterogeneity	
Share of …	Brand Choice	Total*	Brand Choice	Total	Brand Choice	Total	Brand Choice	Total	Only Endogeneity	Endogeneity & Heterogeneity	Category Regression/ Brand Choice	Outside Good
						Effect of Brand A's Price on						
Brand A	−1.194	−1.422	−1.491	−1.689	−1.223	−1.487	−1.930	−2.179	−1.457	−1.669	−1.437	−2.004
	(0.089)	(0.157)	(0.117)	(0.168)	(0.091)	(0.155)	(0.118)	(0.170)	(0.155)	(0.167)	(0.133)	(0.151)
Brand B	0.139	−0.092	0.474	0.273	0.155	0.113	0.735	0.479	0.069	0.102	0.146	0.334
	(0.061)	(0.088)	(0.094)	(0.140)	(0.060)	(0.085)	(0.094)	(0.138)	(0.043)	(0.053)	(0.058)	(0.079)
Brand C	0.550	0.317	0.433	0.232	0.552	0.283	0.470	0.214	0.026	0.049	0.182	0.194
	(0.115)	(0.139)	(0.132)	(0.177)	(0.110)	(0.132)	(0.135)	(0.182)	(0.031)	(0.033)	(0.057)	(0.082)
						Effect of Brand B's Price on						
Brand A	0.206	−0.143	0.708	0.403	0.230	−0.155	1.093	0.724	0.104	0.053	0.216	0.512
	(0.097)	(0.133)	(0.111)	(0.158)	(0.095)	(0.130)	(0.110)	(0.160)	(0.027)	(0.039)	(0.071)	(0.088)
Brand B	−0.434	−0.780	−1.020	−1.320	−0.443	−0.826	−1.381	−1.740	−1.031	−1.230	−1.405	−1.902
	(0.083)	(0.109)	(0.146)	(0.179)	(0.083)	(0.103)	(0.145)	(0.183)	(0.118)	(0.132)	(0.157)	(0.183)
Brand C	0.242	−0.107	0.447	0.142	0.236	−0.149	0.528	0.162	0.049	0.059	0.224	0.365
	(0.119)	(0.156)	(0.139)	(0.188)	(0.118)	(0.155)	(0.139)	(0.186)	(0.024)	(0.028)	(0.073)	(0.081)
						Effect of brand C's Price on						
Brand A	0.992	0.562	0.789	0.414	0.997	0.530	0.851	0.410	0.048	0.089	0.329	0.380
	(0.138)	(0.222)	(0.151)	(0.200)	(0.137)	(0.217)	(0.150)	(0.196)	(0.030)	(0.033)	(0.065)	(0.077)
Brand B	0.294	−0.133	0.544	0.171	0.287	−0.177	0.644	0.204	0.059	0.072	0.274	0.453
	(0.126)	(0.208)	(0.144)	(0.181)	(0.126)	(0.211)	(0.143)	(0.183)	(0.029)	(0.035)	(0.069)	(0.074)
Brand C	−0.793	−1.215	−0.882	−1.251	−0.790	−1.249	−1.004	−1.437	−0.925	−1.035	−1.305	−1.697
	(0.099)	(0.157)	(0.103)	(0.159)	(0.100)	(0.161)	(0.105)	(0.161)	(0.135)	(0.159)	(0.198)	(0.211)

*Brand Choice: Brand Choice Elasticity; Total: Category Sales and Brand Choice Elasticity.

mand elasticities. There are several interesting points to note from Table 3:

(1) None of the cross elasticities are subject to the IIA restriction, even those that come from models that do not account for heterogeneity. This is because of the probit model specification at the individual consumer level. The finding is in contrast with studies such as Berry et al. (1995) and Nevo (2001), where heterogeneity is required to break the IIA restriction due to the logit assumption on brand choices.

(2) As expected, the own elasticity of sales is larger (in magnitude) than the own elasticity of brand choice. However, the cross-elasticities of demand, when positive, are larger for brand choice than for sales. Intuitively, this is because in the case of the own sales elasticity, if the price of a brand increases, consumers can switch to another brand or to not buying at all. Hence, the own elasticity of sales is larger than if consumers are forced to switch to one of the other brands (the brand choice elasticity). With the cross-sales elasticities, an increase in price of brand A implies fewer consumers switch to brands B and C because consumers can also switch to not buying at all. This is not the case for the cross brand choice elasticities.

(3) Comparing own elasticities across model specifications, we find that ignoring either endogeneity or heterogeneity tends to bias the elasticities towards zero. Specifically, the following relationships appear:
  (a) Elasticity (No Endogeneity and No Heterogeneity) < Elasticity (With Endogeneity and No Heterogeneity);
  (b) Elasticity (No Endogeneity and No Heterogeneity) < Elasticity (No Endogeneity and With Heterogeneity).

  Furthermore, we also find that accounting for either endogeneity or heterogeneity does not suffice and it is important to account for both these issues in the estimation. Specifically,
  (a) Elasticity (With Endogeneity and No Heterogeneity) < Elasticity (With Endogeneity and With Heterogeneity);
  (b) Elasticity (No Endogeneity and With Heterogeneity) < Elasticity (With Endogeneity and With Heterogeneity).

(4) Not accounting for endogeneity seems to have a bigger impact than not accounting for unobserved heterogeneity in this category.

(5) Looking at the own sales elasticities from the most general model (accounting for both endogeneity and heterogeneity), we find that brand A is the most price elastic followed by brands B and C. We find that this ordering is preserved when endogeneity is accounted for. However, in the other two cases we find the ordering of brands B and C interchanged even after accounting for the effects of heterogeneity. This further underscores the need to account for these phenomena when estimating the parameters of demand functions.

(6) Examining the cross-sales elasticities from the with endogeneity and heterogeneity model, we find that brand A prices have a bigger impact on the sales of brand B than on brand C. Brand B's price, consistent with that of brand A, has a bigger effect on the sales of that brand than on the sales of brand C. Also, brand A sales seem to be affected the most by brand C's prices with brand B's sales being affected less. These cross-elasticities provide insights into the nature of interbrand price competition in this market.

(7) Note that some of the total sales cross-elasticities have the wrong signs under the two specifications that do not account for endogeneity. The reason for this is that the corresponding brand choice-elasticities are biased toward zero due to not accounting for endogeneity (and, as noted in (4), this has a bigger impact than not accounting for heterogeneity). Recall that the total cross-elasticity is the category sales elasticity + the brand choice cross elasticity. The category elasticity is negatively signed whereas the brand choice cross-elasticity is positively signed. When the latter is biased towards zero, the sum in certain cases turns out to be negative. Note that this is not the case with the brand choice cross-elasticities.

Having discussed the results from the brand choice/category regression specification for the probit model, we turn next to the specification in which an outside good is included in the individual-level choice model to capture the no-purchase behavior of consumers. This obviates the need for a category regression equation. Hence, the model is identical to the brand choice component of the previous specification with an additional alternative. In Table 2, Columns 7–8, we present the parameter estimates and their standard errors for this formulation. Given the relative importance of accounting for endogeneity found with the previous specification, we focus only on the two formulations that account for endogeneity—with and without accounting for heterogeneity.

Note from Table 2 that we now have three brand intercepts—one for each brand. The reason is that we now have four alternatives, the three brands and the outside good, and so three intercepts are identified. The outside good is specified as the base brand in this case. Also note that we have six covariance parameters (of which five are uniquely identified) rather than three as in the previous formulation. The reason is that $\Omega_{j,K-1}$ is now a 3×3 matrix with six unknown parameters in the Cholesky decomposition. The results are largely consistent with those from the category regression/brand choice model. We note once again that the price and promotion parameters are not directly comparable across specifications because of differences in the estimated covariance matrices. We do note however, that the parameters corresponding to the heterogeneity distribution are small and are not significantly different from zero in two of the four cases. Even the standard deviation parameter for price that had an estimated coefficient of 0.242 is only 0.058 in this case. Again, we caution that the numbers are not directly comparable. Nevertheless, they seem to indicate a small effect of heterogeneity in this case. To verify this, we provide in Table 3 (Columns 10–11) the price elasticities from the two specifications. We note the following from these estimates.

(1) Consistent with our previous results, we find that not accounting for the effects of heterogeneity does bias the estimated elasticities towards zero in this case as well.

(2) The relative ordering of the own elasticities is also the same as previously found with brand A being the most price sensitive followed by brands B and C.

(3) The own price elasticities seem to be smaller in this case as compared to the most general model under the category regression/brand choice specification (Column 9 in Table 3). In particular, the own elasticities seem closer to zero by roughly 0.4–0.5 for all three brands. What this implies is that the category elasticities corresponding to this sales specification are smaller than those obtained when category sales were modeled explicitly as a function of category level marketing activities.

(4) Furthermore, the cross-price elasticities are also very small in magnitude, especially compared to those in Columns 2–9. It must be noted that previous studies that have examined the purchase incidence and brand choice decisions of households have also obtained small cross-elasticities relative to own elasticities (Chintagunta 1993).

Taken together, these results imply that the estimated price elasticities are sensitive to the model specification. The choice of specification will come down to a trade-off between wanting a fully structural interpretation of the model versus not having to make assumptions that determine the total size of the category. For example, if one does have data on the entire category's sales, then this information can be exploited in defining the outside good. However, in the absence of such information, the proposed category regression/brand choice model may be preferred.

Model Comparison: Logit Model

Having discussed the results from two different probit specifications, we turn next to the logit model to see whether implications obtained are similar to those obtained for the probit model. Accordingly, in Table 3 (Columns 12–13) we provide the price elasticities obtained from the two logit specifications. The first is a purchase incidence/brand choice model similar to the nested logit model. This is the specification discussed in Chintagunta (1993) except that we allow for the price coefficient to be different from -1. This specification treats the no-purchase option to be distinct from the alternatives in the category under consideration. Hence, even in the absence of heterogeneity the substitution pattern between one of the brands and the outside good is different from that between two brands. The second specification is the category regression/brand choice model, whose direct probit counterpart we have discussed previously. Under both specifications, we account for endogeneity as well as for heterogeneity.

We draw the following inferences from the elasticities in Table 3 (Columns 12 and 13).

(1) Under both specifications, brand A has the highest elasticity, followed by brands B and C in that order. This is consistent with the results obtained from the probit model specifications.

(2) Comparing across specifications, we find that the own elasticities obtained from the purchase incidence/brand choice model are smaller than those obtained from the category regression/brand choice model. Note that the former specification requires an assumption on category consumption much like the probit model with the outside good. Furthermore, the magnitude of difference in elasticities is roughly comparable to the differences from the corresponding probit models.

(3) Comparing the elasticities from the category regression/brand probit and logit models in Table 3, we find that these elasticities are quite comparable in their magnitudes. The own elasticities for the three brands under the probit specification are -2.179, -1.740, and -1.437. The corresponding elasticities from the logit model are -2.004, -1.902, and -1.679. It appears from these numbers that the logit elasticities vary over a smaller range than the probit elasticities. In other words, optimal margins for the manufacturers under the probit specification will lead to a wider range in margins than under the logit specification. Performing this computation, we find the margin ((price $-$ cost)/price) for the three brands under the probit specification to be 46%, 57%, and 69%. Under the logit specification, we obtain 50%, 53%, and 60%. Similarly, the cross-elasticities range from 0.194 to 0.512 under the logit specification, whereas they range from 0.162 to 0.724 under the probit model.

(4) A comparison of the elasticities from the probit outside good model from Table 3 (Column 11) with the logit purchase incidence/brand choice model in Table 3 (Column 12) reveals a pattern similar to that of the comparison described above. However, in this case it appears that the logit own price elasticities across the three brands are very close to one another, ranging only from -1.305 for brand C to -1.437 for brand A. Hence, it appears that the logit assumption on brand choice probabilities may be restricting the range of elasticities estimated from the data. This provides further motivation for using the probit model to characterize demand when studying competitive interactions among firms.

To summarize, the model comparison results indicate that while the elasticities from the logit and probit specifications are roughly comparable, there are some differences that exist. As these differences have implications for optimal pricing behavior, they are of interest to researchers studying competitive behavior at the firm level. Given the more flexible nature of the choice model under the probit specification, one can, for these data, conclude that this is a more appropriate specification. This is notwithstanding the flexibility imparted to the logit model by the distribution of heterogeneity imposed. We also carried out a predictive validation exercise on four holdout weeks for the logit and probit outside good models. Note that predictions with such models require us to also integrate over the distribution of the unobserved attribute, as we do not observe these terms for the holdout data. We use the empirical distribution for the purpose and make predictions at each of the 104 unobserved attribute values from the estimation sample. The average share for each brand across the 104 values is computed for each hold out week. The mean absolute percentage error from the logit model is 27% and that for the probit model is 23%.

Model Comparison: Log–Log Regression Model

We also estimated the parameters from a log–log regression model that is most comparable to our model specification (i.e., using the same set of variables). We also estimated the linear and semilog regression models, as some of these specifications are more appropriate for pricing purposes. The price elasticities from the log–log model are in Table 4, with those from the other specifications being substantively sim-

Table 4 Price Elasticities from Log–Log Regression Model

Sales of/Price of →	Brand A	Brand B	Brand C
Brand A	−0.129	−1.198	1.232
	(0.599)	(0.489)	(0.813)
Brand B	−1.439	1.799	−0.836
	(0.568)	(0.601)	(0.885)
Brand C	2.205	−0.347	−2.633
	(0.532)	(0.545)	(0.821)

ilar. We find from Table 4 that the own-price elasticity for brand B has the incorrect sign, while that for brand A is not significantly different from zero. Of the cross-elasticities, two are positive, two negative, and two are not significantly different from zero. As noted in the introduction, using these elasticities as the basis for strategic pricing decisions can be problematic.

Conclusions

In this paper, we have proposed the probit model as an alternative to the logit model to specify the aggregate demand functions of firms competing in oligopoly markets. The primary benefit that accrues from using the probit model is an avoidance of the IIA property at the individual consumer level that enables us to distinguish between the effects of IIA violations and the effects of heterogeneity at the aggregate level. In the estimation of the model parameters, we account for two critical issues that have received recent attention in the marketing literature. These are endogeneity of marketing variables and heterogeneity across consumers. The endogeneity problem arises because of unobserved factors that are firm- and time period-specific (but invariant across consumers) that could be correlated with price. Consumer heterogeneity is accounted for by assuming that brand preferences and price sensitivities vary across consumers following a parametric distribution. The individual level choice probabilities are aggregated across heterogeneous consumers, and the aggregated demand function is taken to the data.

Our results indicate that both endogeneity as well as heterogeneity need to be accounted for even after allowing for a non-IIA specification at the individual consumer level. We also find that ignoring endogeneity has a bigger impact on the estimated price elasticities than ignoring the effects of heterogeneity. A comparison of the elasticities obtained from the probit model with those from the corresponding logit specification indicates that while the elasticities appear to be comparable in magnitude, there is one key difference. We find that the *range* of elasticities obtained from the probit model across brands is larger than that obtained from the logit. This finding could stem in part from the probit model, allowing for different error variances across brands.

In addition to specifying a probit model and providing comparisons with the logit model, the paper also addresses the issue of the specification of the ''outside good'' that arises when using discrete choice models to specify demand. We propose a simple alternative to this specification by decomposing the demand for a brand into a category demand equation and a conditional brand choice share equation. We provide a comparison of results from this specification to those from the outside good specification and find that estimated elasticities are sensitive to the specification used.

One of the key limitations of the proposed model is that while it accounts for the purchase incidence and brand choice decisions of households, it does not account for differences across consumers in their purchase quantities. The model and analysis are best suited for product categories in which consumers typically make only single-unit purchases. Another limitation is more practical in nature. While recent advances have been made in computing probit probabilities, it could nevertheless be a challenge to do so when the number of alternatives is large.

In summary, this study has proposed a probit demand model as an alternative to the logit model that can be used as a basis to investigate competitive interactions among firms in a product market. We examine the sensitivity of the estimated price elasticities to the specification of the no-purchase alternative in these models. The estimation of the model parameters accounts for endogeneity and heterogeneity. Our model results obtained from the analysis of the shampoo product category indicate that the proposed specification is a promising alternative to existing methods used for the purpose.

Acknowledgments
The author thanks the editor, the area editor, and two anonymous reviewers for their comments and suggestions. He is grateful to J. P. Dube for useful discussions. The author thanks the Kilts Center at the University of Chicago for partial funding of this research.

References

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

Berry, S. T., M. Carnall, P. T. Spiller. 1998. Airline hubs: Costs and markups and the implications of consumer heterogeneity. NBER Working Paper 5561.

——, J. Levinsohn, A. Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* **60** (4) 841–890.

Besanko, D., S. Gupta, D. Jain. 1998. Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Sci.* **44** (11) 1533–1547.

Chintagunta, P. K. 1993. Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Sci.* **12** (2) 184–208.

——, B. E. Honore. 1996. Investigating the effects of marketing variables and unobserved heterogeneity in a multinomial probit model. *IJRM* **13** 1–15.

Christen, M., S. Gupta, J. Porter, R. Staelin, D. R. Wittink. 1997. Using market-level data to understand promotion effects in a nonlinear model. *J. Marketing Res.* **34** 322–334.

Currim, I. S. 1982. Predictive testing of consumer choice models not subject to independence of irrelevant alternatives. *J. Marketing Res.* **19** (2) 208–222.

Hajivassiliou, V. A., D. L. McFadden, P. A. Ruud. 1996. Simulation of multivariate normal rectangle probabilities and their derivatives: Theoretical and computational results. *J. Econometrics* **72** (1–2) 85–134.

Jain, D. C., N. J. Vilcassim, P. K. Chintagunta. 1994. A random-coefficients logit brand-choice model applied to panel data. *J. Bus. Econom. Statist.* **12** (July) 317–328.

Kadiyali, V. 1996. Entry, its deterrence, and its accommodation: A study of the U.S. photographic film industry. *Rand J. Econom.* **27** (3) 452–478.

Kim, B-D., R. C. Blattberg, P. E. Rossi. 1995. Modeling the distribution of price sensitivity and implications for optimal retail pricing. *J. Bus. Econom. Statist.* **13** (3) 291–303.

Nevo, A. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica* **69** (2) 307–342.

Sudhir, K. 2001. Competitive pricing behavior in the auto market: A structural analysis. *Marketing Sci.* **20** (1) 42–60.

Villas-Boas, M., R. Winer. 1999. Endogeneity and brand choice models. *Management Sci.* **45** (10).

This paper was received December 20, 1999, and was with the author 9 months for 4 revisions; processed by Greg Allenby.