Endogeneity in Brand Choice Models*

J. Miguel Villas-Boas**

Russell S. Winer***

October, 1995

* We are grateful to Daniel McFadden for long and insightful conversations. We are also grateful to Tülin Erdem, Michael Keane, Don Lehmann, and four anonymous reviewers for comments on an earlier version of this paper. All remaining errors are our responsibility alone. This work is supported by the NSF grant SES-9309812. Research assistance by Mark Stiving and Ying Zhao is gratefully appreciated.

** Assistant Professor, Haas School of Business, University of California at Berkeley, Berkeley CA 94720.

*** J. Gary Shansby Professor of Marketing Strategy, Haas School of Business, University of California at Berkeley, Berkeley CA 94720.

Endogeneity in Brand Choice Models

ABSTRACT

Applications of random utility models to scanner data have been widely presented in marketing for the last 20 years. One particular problem of these applications is that they have ignored possible correlations between the independent variables in the deterministic component of utility (price, promotion, etc.) and the stochastic component or error term. Marketing-mix variables, such as price, not only affect brand purchasing probabilities but are themselves endogenously set by marketing managers. This work tests whether these endogeneity problems are important enough to warrant consideration when estimating random utility models with scanner panel data. Our results show that not accounting for endogeneity may result in a substantial bias in the parameter estimates.

1. Introduction

Over the past 20 years, there has been a significant increase in the research literature on the topic of modeling the choice process of economic agents or consumers. Motivated by McFadden's (1973) work on the conditional multinomial logit model, there has been particular interest in what are referred to as random utility models where the decision maker faces a choice set for which the utility of each alternative is a random variable. This utility is usually specified to have a deterministic component which is a function of observable variables and a random component for which a variety of parametric assumptions have been made. In particular, random utility models, mainly logit and probit models, have been often applied to household purchasing behavior using electronic scanner data (e.g. Guadagni and Little 1983).

One particular problem in how marketing researchers have applied these random utility models to scanner data is that they have ignored the fact that the exogenous deterministic components of utility such as price, promotion, advertising, etc. are themselves endogenous. Marketing managers set these marketing-mix variables based on market information which may be in part unobservable to the researcher but which nevertheless affects consumer choice. This would create a situation where the marketing-mix variables could be correlated with the residuals in the latent utilities.[1] Failure to account for this endogeneity in the deterministic components of utility has the potential to bias the parameter estimates of the marketing-mix variables. If this happens, both major uses of these models, for diagnostic and optimization purposes, could produce misleading results, seriously affecting the outcomes of the marketing decisions.[2] This problem of endogeneity is not, of course, a new topic in the marketing literature as Bass (1969) discussed the problem in the context of the simultaneity issue between advertising and sales.

The justification researchers have presented for not accounting for these endogeneity problems in brand choice models with scanner data has to do with the fact that, in a certain period, the marketing-mix variables are common across all consumers. If the utilities' error terms are independent across individuals, it is unlikely that the marketing-mix variables (common across all individuals) are highly correlated with the error terms. However, in this paper, we question this independence assumption and show that this assumption may not hold which may result in a correlation between the error terms and the marketing-mix variables.

Why may the error terms in the latent utilities not be independent across individuals? One possibility is that there are idiosyncratic word-of-mouth or fashion effects. Alternatively, there are always market phenomena that affect all households and are not observable by the researcher (and, therefore, not included in the choice model), but that the marketing managers use in their decisions. These are referred in what follows as common demand "shocks". One example of such a phenomenon is significant activity in a competing product class. A significant price increase in coffee due to a freeze in South America may produce an increase in demand for a substitute like colas and therefore affect both the cola brand choice (more stockouts due to higher demand) and the decision-making environment (higher prices than one might predict for some brands). A second example could be a local or regional supermarket promotional war outside of the category being modeled but which affects brand choice in that category.

The major purpose of this work is then to test whether these endogeneity problems are sufficiently important to warrant consideration when estimating random utility models with scanner panel data. In particular, we focus on testing endogeneity of the pricing decision. While testing for endogeneity, we also obtain unbiased estimates for the random utility model under the assumption that the endogeneity (firms' decision-making) model we use is the true one. Several theoretical tests for endogeneity in models of limited dependent variables have been presented by, among others, Heckman (1978), Newey (1985, 1987), and Rivers and Vuong (1988). Implementations of these tests have not been very frequent (one example is Heckman 1979) and we do not know of any applications in the context of brand choice models (where an explanatory variable, like price, has the same value for several individuals). Other related work is Berry, Levinsohn, and Pakes (1995) who only look at aggregate data, Goldberg (1995) who does not account for endogeneity, and Berry, Carnall, and Spiller (1995).

The tests are performed in two frequently purchased product categories: yogurt and ketchup. We find that, subject to the assumptions of our model, the bias by ignoring endogeneity can be substantial in that there are important changes in the measures of price sensitivity and the individual brand alternative parameters. This is in contrast with the small impact on the average effect on the price parameter when one allows for heterogeneity (Allenby and Rossi 1991) or for the no-purchase option (Gupta 1988). We also show that the base model results are robust to different specifications of the error structure, to heterogeneity in the consumer pool, and to different specifications of the firms' decision making.

The paper is organized as follows. In section 2 we present the results of a base model that summarizes the main message of the paper. This model has a very simple error structure, considers endogeneity on the additive term of the random utility model, and does not account for heterogeneity (though it includes a loyalty variable). Section 3 presents the extension to more general error structures, Section 4 considers endogeneity on the slope, and Section 5 accounts for heterogeneity in the consumer pool. Section 6 relaxes the pricing rule being used, and Section 7 concludes and discusses directions for future research.

2. A Base Model with Endogeneity

This section considers a base random utility model with endogeneity. Consumers are assumed to buy one product from a choice set S with J alternatives. A consumer derives utility only from buying one product in that choice set.[3] The consumer buys the product for which the perceived utility minus the price (which we call the indirect utility) is the greatest. In this section, utility has a deterministic component and a random component. The deterministic component has a fixed term, a signaling effect term (with the variables feature and display; see, for example, Milgrom and Roberts 1986), and a loyalty term as a way of accounting for heterogeneity among consumers (we model heterogeneity directly in Section 5 below). In this base model, the loyalty variable equals one if the consumer purchased that product the last time she purchased any product in this category, and is set to zero otherwise.[4]

Given that all that matters in the decision of which product to choose is which one yields the highest indirect utility, the form described above is unique up to a monotonic transformation. We can then write the indirect utility function as

$$U_{ijt} \;=\; X'_{ijt}\beta \;+\; \varepsilon^a_{ijt} \;+\; \varepsilon^b_{jt} \tag{1}$$

where i indexes households, j represents the brands, and t represents the week in which household i purchased a product from the choice set C (the set of consumers that made a purchase in week t is referred to as I_t; the product chosen by consumer i in week t is referred to as j_{it}). X_{ijt} is a vector with the following variables: dummy variables for the number of alternatives minus one, price, display, feature, and loyalty. β is the vector of parameters to be estimated. U_{ijt} is a latent variable; the researcher observes the product being chosen, which we call d_{it}. The error term ε^a_{ijt} is assumed to be Gumbel distributed with parameters $(0,\,\theta)$.[5] The error term ε^b_{jt} is assumed to be normally distributed with mean zero and variance $\sigma^2_{\varepsilon^b}$. Furthermore, it is assumed that all the elements of the set $\{\varepsilon^a_{ijt}, \varepsilon^b_{jt} \forall i, j, t\}$ are independent from each other (in some of the next sections we allow for more general error structures). The scale of the indirect utility is set by $\sigma^2_{\varepsilon^b} + \frac{\pi^2}{6\theta^2} = \frac{\pi^2}{6}$.

The main difference from the usual models is the existence of ε^b_{jt} which represents the common demand shocks of the type described in the Introduction.[6] If $\sigma^2_{\varepsilon^b} = 0$, this model reduces to the usual models. Testing for common demand shocks is testing for the null hypothesis that $\sigma^2_{\varepsilon^b} = 0$.

In this section we introduce endogeneity in prices. In order to do this, we assume that the firms set prices according to the rule

$$P_{jt} = \alpha_{j0} + \alpha_{j1} P_{j,t-1} + \eta_{jt} \tag{2}$$

where P_{jt} is the price set for brand j in period t, and α_{j0} and α_{j1} are parameters (the pricing rule is relaxed in Section 6). This approach can also be seen as using lagged prices as instruments for current prices. This is a "limited information" approach because we do not include in the analysis any restrictions of the parameters across equations. The main advantage of this approach is that specification problems in the pricing equation do not extend to the parameters of the utility equation – the main objective of the analysis.[7]

This pricing rule can be derived from profit maximization if the costs are correlated through time and the firms observe only a certain part of the demand shocks. Alternatively, it can be simply seen as a rule justified by bounded rationality on the part of the firms' decision-makers. Note that this pricing rule is sufficiently general that the current price may result from both retailers and manufacturers activities (or just retailers, or just manufacturers). The error η_{jt} consists of unobservables (to the researcher) which reflect shocks on costs (for example, productivity, input prices) and on demand (part of ε^b_{jt}). Because η_{jt} consists in part of demand shocks, it is very likely that η_{jt} is correlated with ε^b_{jt}. This is the dimension of the endogeneity problem that is studied in this section. We assume that η_{jt} is normally distributed, with mean zero, and variance σ^2_η, and that $E[\eta_{jt}\varepsilon^b_{jt}] = \rho\sigma_{\varepsilon^b}\sigma_\eta$. Testing for the endogeneity problem to be significant is testing for the null hypothesis that $\rho = 0$.

It is also assumed in this section that

$$E[\eta_{jt}\eta_{j't'}] = 0 \;\; if\; j \neq j' \; or \; t \neq t' \tag{3}$$

$$E[\eta_{jt}\varepsilon^b_{j't'}] = 0 \;\; if\; j \neq j' \; or \; t \neq t' \tag{4}$$

and that all the elements of the set $\{\varepsilon^a_{ijt}, \eta_{jt} \forall i, j, t\}$ are independent from each other.

Assuming η_{jt} to be independent across j requires, given existing game-theoretic models of interaction among firms, that the firms are only able to observe each of their own demand shocks and not the competitors' demand shocks. This is relaxed below.[8]

This model can be estimated using maximum likelihood. The likelihood function can be written as

$$L(\beta, \alpha_0, \alpha_1, \sigma^2_{\varepsilon^b}, \sigma^2_\eta, \rho) = \prod_{t=1}^{T}\prod_{h=1}^{J} f(\eta_{ht}; \sigma^2_\eta; \alpha_0, \alpha_1) g(d_{it} \; \forall i \in I_t \mid \eta_{jt} \; \forall j) \tag{5}$$

where α_0 and α_1 are vectors with generic elements, respectively, α_{j0} and α_{j1}, $f(\eta_{ht}; \sigma^2_\eta; \alpha_0, \alpha_1)$ is the density function of a normal random variable with mean zero and variance σ^2_η, and $g(d_{it} \; \forall i \in$

$I_t \mid \eta_{jt}\ \forall j)$ is the probability of observing choices $d_{it}\ \forall i \in I_t$ in week t given $\eta_{jt}\ \forall j \in J$. Finally, $g(d_{it}\ \forall i \in I_t \mid \eta_{jt}\ \forall j)$, the conditional probability of observing the purchases in week t given the shocks in the pricing equation, can be written as

$$g(d_{it}\ \forall i \in I_t \mid \eta_{jt}\ \forall j) = \int \int \cdots \int \prod_{i \in I_t} \frac{e^{\theta(X'_{ij_{it}t}\beta + \varepsilon^b_{j_{it}t})}}{\sum_{j=1}^{J} e^{\theta(X'_{ijt}\beta + \varepsilon^b_{jt})}} \prod_{k=1}^{J} f(\varepsilon^b_{kt} \mid \eta_{kt}; \rho, \sigma^2_{\varepsilon^b}, \sigma^2_\eta) d\varepsilon^b_{kt} \qquad (6)$$

where $f(\varepsilon^b_{kt} \mid \eta_{kt}; \rho, \sigma^2_{\varepsilon^b}, \sigma^2_\eta)$ is the conditional density distribution of the normally distributed random variable ε^b_{kt} on the normally distributed random variable η_{kt}, both with zero mean, with correlation ρ, and variances, respectively, $\sigma^2_{\varepsilon^b}$ and σ^2_η.

If endogeneity is important (i.e., $\rho \neq 0$) one might expect prices to be positively correlated with the residuals of the latent utilities (i.e., with greater demand firms set higher prices) which results in $\rho > 0$. Furthermore, because of this positive correlation, we expect that the failure to account for endogeneity in the choice model results in the price effect being underestimated.

This model was estimated using scanner panel data from both the yogurt and the ketchup markets. In the yogurt market, we restricted our attention to C={Dannon,Yoplait,Private Label} (Dannon and the Private Label were in 8 oz. packages; Yoplait was in a 6 oz. package); these brands account for 77% of the market. There are 3,513 purchase occasions on C in the data being used. The purchases occur during 137 weeks. We restrict our attention to only the largest store in order to simplify the analysis.[9] In the ketchup market we restricted our attention to C={DelMonte,Hunts,Heinz} for the 32 oz. package, which accounts for 57% of the market.[10] There are 1256 purchase occasions on C in the data being used. The purchase occasions occur during 129 weeks. We restrict our attention to only one store of the three largest ones.[11] The results are presented in Table 1.

These results clearly show the importance of considering the endogeneity of the pricing decision. First, there are important common shocks across individuals as the hypothesis $\sigma_{\varepsilon^b} = 0$ is rejected for both product categories. Note that the value of the chi-square statistic on this restriction is 51.6 for the yogurt market and 33.0 for the ketchup market[12] (the critical value at the 5% significance level is 3.84). Second, there is significant endogeneity of the price variable as the hypothesis $\rho = 0$ is rejected for both product categories. Note that the chi-square statistic is 36.8 for the yogurt market and 5.0 for the ketchup market (the critical value at the 5% significance level is also 3.84). Third, the parameter that estimates the impact of price on choice seems to be clearly underestimated (expected direction of the bias) if endogeneity is not taken into account. For the yogurt market the price effect is underestimated by 57%; for the ketchup market, it is underestimated by 11%. Note also that the

Table 1
Results from the Base Case

	Complete Model	$\rho = 0$	$\rho = 0, \sigma_{\varepsilon^b} = 0$
Yogurt Market			
Dannon	5.67	3.48	3.32
	(0.07)	(0.08)	(0.12)
Yoplait	5.75	3.52	3.34
	(0.10)	(0.08)	(0.13)
Price	-21.74	-14.22	-13.83
	(0.20)	(0.31)	(0.45)
Display	0.33	0.51	0.54
	(0.03)	(0.04)	(0.09)
Feature	-0.015	-0.147	-0.069
	(0.02)	(0.02)	(0.07)
Loyalty	1.63	1.75	1.79
	(0.03)	(0.04)	(0.04)
σ_η	0.039	0.038	0.038
	(0.001)	(0.002)	(0.002)
σ_{ε^b}	0.61	0.40	
	(0.04)	(0.09)	
ρ	0.78		
	(0.05)		
LL	-1086.2	-1104.6	-1130.4
Ketchup Market			
DelMonte	-2.89	-2.54	-2.46
	(0.13)	(0.16)	(0.26)
Hunts	-2.69	-2.43	-2.40
	(0.13)	(0.15)	(0.25)
Price	-6.21	-5.61	-5.58
	(0.21)	(0.31)	(0.56)
Display	1.22	1.16	1.14
	(0.11)	(0.10)	(0.07)
Feature	0.80	0.76	0.83
	(0.12)	(0.10)	(0.10)
Loyalty	1.24	1.25	1.25
	(0.07)	(0.07)	(0.08)
σ_η	0.056	0.056	0.056
	(0.001)	(0.001)	(0.001)
σ_{ε^b}	0.56	0.50	
	(0.08)	(0.09)	
ρ	0.30		
	(0.09)		
LL	97.6	95.1	78.6

LL is the value of the log likelihood minus the fixed terms. The standard errors are in parentheses (this will also be the case for all the tables below).

correlation between ε^b_{jt} and η_{jt}, ρ, is positive as expected (note that it is quite large for the yogurt market). The common demand shocks are 23% of the total demand shocks for the yogurt market, and 19% for the ketchup market. Finally, note that there is a pattern in the change of the estimates of the parameters of the latent utility model when the endogeneity of the pricing decision is allowed for: the brand-specific dummies increase in absolute value, and the effect of the other marketing-mix variables and of loyalty decrease (for most of the coefficients).

These results present the main message of the paper: accounting for the endogeneity of a marketing-mix variable in a brand choice model can result in a significant change in the parameter estimates, and therefore, in the resulting optimal resource allocation. In the next sections we are going to show that these results are robust to more general error structures, to a slope specification of the endogeneity, to heterogeneity in the consumer pool, and to relaxations of the pricing rule.

3. More General Error Structures

In this section we consider the case of more general error structures. In particular, we consider structures where the independence of irrelevant alternatives property does not hold and where there might be non-zero correlations across j between the elements of the set $\{\varepsilon^b_{jt}, \eta_{jt} \forall j, t\}$.

In order to allow the independence of irrelevant alternatives property not to hold, we assume ε^a_{ijt} to be normally distributed instead of Gumbel. As is well-known, this creates major computational problems when using maximum likelihood estimation (in this particular case, we have multiple integrals of the fourth order). As a result, we use the method of simulated maximum likelihood (SML).[13] For some simulators, SML is equivalent to the method of simulated scores (see Keane 1993 and Hajivassiliou and McFadden 1993, and Gourieroux and Monfort 1993 for a discussion of the properties of these estimators).[14]

Then

$$g(d_{it}\ \forall i \in I_t \mid \eta_{jt}\ \forall j) = \int\int\cdots\int \prod_{i\in I_t} Prob(j_{it}; i, t, \varepsilon^b_{jt}\forall j, t) \prod_{k=1}^{J} f(\varepsilon^b_{kt} \mid \eta_{kt}; \rho, \sigma^2_{\varepsilon^b}, \sigma^2_\eta) d\varepsilon^b_{kt} \qquad (7)$$

where

$$Prob(j_{it}; i, t, \varepsilon^b_{jt}\forall j, t) \quad = \quad \tilde{\Phi}[(X'_{ij_{it}t} - X'_{i1t})\beta + \varepsilon^b_{j_{it}t} - \varepsilon^b_{1t}, \cdots, (X'_{ij_{it}t} - X'_{i(j-1)t})\beta +$$

$$\varepsilon^b_{j_{it}t} - \varepsilon^b_{(j-1)t}, (X'_{ij_{it}t} - X'_{i(j+1)t})\beta +$$
$$\varepsilon^b_{j_{it}t} - \varepsilon^b_{(j+1)t}, \cdots, (X'_{ij_{it}t} - X'_{iJt})\beta + \varepsilon^b_{j_{it}t} - \varepsilon^b_{Jt}] \tag{8}$$

where $\tilde{\Phi}[.]$ is the cumulative normal distribution of dimension $J-1$ of a vector with mean zero and variance-covariance the variance-covariance of the vector $(\varepsilon^a_{i1t} - \varepsilon^a_{ij_{it}t}, \cdots, \varepsilon^a_{i(j-1)t} - \varepsilon^a_{ij_{it}t}, \varepsilon^a_{i(j+1)t} - \varepsilon^a_{ij_{it}t}, \cdots, \varepsilon^a_{iJt} - \varepsilon^a_{ij_{it}t})$. This can be obtained, for example, from the variance-covariance of the vector $(\varepsilon^a_{i2t} - \varepsilon^a_{i1t}, \varepsilon^a_{i3t} - \varepsilon^a_{i1t}, \cdots, \varepsilon^a_{iJt} - \varepsilon^a_{i1t})$. The elements of this matrix are parameters of the likelihood function. In this particular case, because $J = 3$, the parameters are

$$V = VAR\begin{bmatrix} \varepsilon^a_{i2t} - \varepsilon^a_{i1t} \\ \varepsilon^a_{i3t} - \varepsilon^a_{i1t} \end{bmatrix} = \begin{bmatrix} \sigma^2_{\varepsilon^a_{2-1}} & \rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}\sigma_{\varepsilon^a_{2-1}}\sigma_{\varepsilon^a_{3-1}} \\ & \sigma^2_{\varepsilon^a_{3-1}} \end{bmatrix} \tag{9}$$

where now the indirect utility is scaled by $\sigma^2_{\varepsilon^a_{2-1}} + 2\sigma^2_{\varepsilon^b} = 1$.[15]

Using SML in this model requires simulation of both $\tilde{\Phi}[.]$ (draws of ε^a_{ijt}) and $g(.)$ given $\tilde{\Phi}[.]$ (draws of ε^b_{jt}). The simulation of $\tilde{\Phi}[.]$ was performed using the GHK simulator (Geweke-Hajivassiliou-Keane; see, for example, Hajivassiliou, McFadden and Ruud 1992 for a description of the simulator). Because it is an expected value of a smooth function, we simulated $g(.)$ by taking the mean across draws of ε^b_{jt}.

In order to check the accuracy of the simulation of $g(.)$ alone, we first simulated $g(.)$ for the model of the previous Section (i.e., keeping ε^a_{ijt} with a Gumbel distribution). The results, not presented here (but available upon request from the authors), are very similar to the results presented in the previous Section.

The results for the case in which ε^a_{ijt} is normally distributed (still assuming η_{jt}, and ε^b_{jt} to be identically distributed and independent across j) are presented in Table 2 for the yogurt and the ketchup markets. Columns 4 through 6 present the results that are directly comparable (up to a constant) to the results of the logit case, i.e., V is the identity matrix multiplied by a constant. Columns 1 through 3 present the general case for V, where the independence of irrelevant alternatives property is relaxed, and the elements in the diagonal of V do not need to be all equal.

These results confirm the results obtained in the base case and show their robustness to allowing the independence of irrelevant alternatives not to hold. They show clearly the importance of considering the endogeneity of the pricing decision: (i) there are important common shocks across individuals (the hypothesis $\sigma_{\varepsilon^b} = 0$ is rejected), (ii) there is significant endogeneity of the price variable (the hypothesis $\rho = 0$ is rejected), and (iii) the parameters that estimate the impact of price on choice are underestimated if endogeneity is not taken into account. For the yogurt market the price effect is underestimated by 65%; for the ketchup market, it is underestimated by 13% (in comparison

Table 2
Simulation of $g(.)$ and $\tilde{\Phi}\,[.]$

	Complete Model	$\rho = 0$	$\rho = 0, \sigma_{\varepsilon^b} = 0$	$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}} = 0$ $\sigma_{\varepsilon^a_{2-1}} = \sigma_{\varepsilon^a_{3-1}}$	$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}} = 0$ $\sigma_{\varepsilon^a_{2-1}} = \sigma_{\varepsilon^a_{3-1}}$ $\rho = 0$	$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}} = 0$ $\sigma_{\varepsilon^a_{2-1}} = \sigma_{\varepsilon^a_{3-1}}$ $\rho = 0, \sigma_{\varepsilon^b} = 0$
Yogurt Market						
Dannon	2.69	1.60	1.49	3.90	2.05	1.87
	(0.28)	(0.11)	(0.08)	(0.23)	(0.10)	(0.06)
Yoplait	2.79	1.70	1.56	3.73	1.78	1.60
	(0.29)	(0.11)	(0.09)	(0.26)	(0.11)	(0.07)
Price	-10.55	-6.81	-6.41	-14.24	-7.84	-7.27
	(0.96)	(0.38)	(0.30)	(0.78)	(0.36)	(0.21)
Display	0.20	0.23	0.24	0.23	0.31	0.33
	(0.07)	(0.08)	(0.05)	(0.07)	(0.10)	(0.06)
Feature	0.027	-0.033	-0.016	0.025	-0.065	-0.039
	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)	(0.04)
Loyalty	0.80	0.83	0.85	0.83	1.00	1.01
	(0.04)	(0.04)	(0.04)	(0.05)	(0.02)	(0.02)
σ_η	0.039	0.038	0.038	0.040	0.038	0.038
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
σ_{ε^b}	0.29	0.16		0.42	0.20	
	(0.04)	(0.02)		(0.04)	(0.02)	
ρ	0.84			0.94		
	(0.05)			(0.02)		
$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}$	0.68	0.71	0.67			
	(0.03)	(0.03)	(0.03)			
$\sigma_{\varepsilon^a_{3-1}}$	0.83	0.90	0.90	0.80	0.96	1.00
	(0.04)	(0.04)	(0.04)			
LL	-1093.8	-1110.4	-1133.6	-1167.3	-1198.8	-1216.1
Ketchup Market						
DelMonte	-1.29	-1.18	-1.10	-1.31	-1.17	-1.03
	(0.13)	(0.12)	(0.11)	(0.17)	(0.16)	(0.12)
Hunts	-1.29	-1.20	-1.16	-1.26	-1.15	-1.07
	(0.12)	(0.12)	(0.10)	(0.16)	(0.15)	(0.12)
Price	-2.48	-2.25	-2.19	-3.24	-2.96	-2.79
	(0.28)	(0.27)	(0.24)	(0.34)	(0.32)	(0.26)
Display	0.52	0.53	0.50	0.69	0.70	0.66
	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.04)
Feature	0.41	0.36	0.42	0.41	0.37	0.45
	(0.08)	(0.07)	(0.05)	(0.10)	(0.09)	(0.05)
Loyalty	0.59	0.59	0.60	0.66	0.67	0.68
	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)
σ_η	0.056	0.056	0.056	0.056	0.056	0.056
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
σ_{ε^b}	0.23	0.22		0.29	0.27	
	(0.03)	(0.03)		(0.04)	(0.04)	
ρ	0.27			0.23		
	(0.12)			(0.14)		
$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}$	0.57	0.53	0.45			
	(0.10)	(0.10)	(0.10)			
$\sigma_{\varepsilon^a_{3-1}}$	0.50	0.51	0.57	0.92	0.93	1.00
	(0.05)	(0.05)	(0.05)			
LL	111.9	109.4	94.9	88.5	86.4	71.8

LL is the value of the log likelihood minus the fixed terms. The number of draws is 10 for the GHK and 10 for ε^b_{jt}.

to the case where endogeneity and common errors are not considered, i.e., column 1 versus column 3). Allowing for a more general error structure is also clearly important. The common demand shocks are on average 36% of the total demand shocks in the yogurt market and 19% in the ketchup market.

In order to check the impact of the number of draws on the previous results, we derived the results of the three left columns of Table 2 with a greater number of draws for the GHK (20) and for ε^b_{jt} (100). These results are presented in Table 3.

A comparison of the results of Table 3 to the first three columns of Table 2 suggests that the bias introduced by SML is not substantial. Furthermore, the SML method is underestimating the likelihood function when we relax the common errors and correlation restrictions, i.e., the SML method is underestimating the impact of endogeneity.

Finally, one can generalize the error structure even further by allowing η_{jt} and ε_{jt} to be distributed differently and correlated across the brands. In particular, for the cases we are analyzing (i.e., $J = 3$) we have the following parameters:

$$V_\eta = VAR\begin{bmatrix}\eta_{1t}\\ \eta_{2t}\\ \eta_{3t}\end{bmatrix} = \begin{bmatrix}\sigma^2_{\eta_1} & \rho_{\eta_{12}}\sigma_{\eta_1}\sigma_{\eta_2} & \rho_{\eta_{13}}\sigma_{\eta_1}\sigma_{\eta_3}\\ & \sigma^2_{\eta_2} & \rho_{\eta_{23}}\sigma_{\eta_2}\sigma_{\eta_3}\\ & & \sigma^2_{\eta_3}\end{bmatrix} \tag{10}$$

$$V_{\varepsilon^b} = VAR\begin{bmatrix}\varepsilon^b_{1t}\\ \varepsilon^b_{2t}\\ \varepsilon^b_{3t}\end{bmatrix} = \begin{bmatrix}\sigma^2_{\varepsilon^b_1} & \rho_{\varepsilon^b_{12}}\sigma_{\varepsilon^b_1}\sigma_{\varepsilon^b_2} & \rho_{\varepsilon^b_{13}}\sigma_{\varepsilon^b_1}\sigma_{\varepsilon^b_3}\\ & \sigma^2_{\varepsilon^b_2} & \rho_{\varepsilon^b_{23}}\sigma_{\varepsilon^b_2}\sigma_{\varepsilon^b_3}\\ & & \sigma^2_{\varepsilon^b_3}\end{bmatrix} \tag{11}$$

$$V_{\varepsilon^b\eta} = COV\begin{bmatrix}\varepsilon^b_{1t}\\ \varepsilon^b_{2t}\\ \varepsilon^b_{3t}\end{bmatrix}\begin{bmatrix}\eta_{1t}\\ \eta_{2t}\\ \eta_{3t}\end{bmatrix} = \begin{bmatrix}\rho_{\varepsilon^b_1\eta_1}\sigma_{\varepsilon^b_1}\sigma_{\eta_1} & \rho_{\varepsilon^b_1\eta_2}\sigma_{\varepsilon^b_1}\sigma_{\eta_2} & \rho_{\varepsilon^b_1\eta_3}\sigma_{\varepsilon^b_1}\sigma_{\eta_3}\\ \rho_{\varepsilon^b_2\eta_1}\sigma_{\varepsilon^b_2}\sigma_{\eta_1} & \rho_{\varepsilon^b_2\eta_2}\sigma_{\varepsilon^b_2}\sigma_{\eta_2} & \rho_{\varepsilon^b_2\eta_3}\sigma_{\varepsilon^b_2}\sigma_{\eta_3}\\ \rho_{\varepsilon^b_3\eta_1}\sigma_{\varepsilon^b_3}\sigma_{\eta_1} & \rho_{\varepsilon^b_3\eta_2}\sigma_{\varepsilon^b_3}\sigma_{\eta_2} & \rho_{\varepsilon^b_3\eta_3}\sigma_{\varepsilon^b_3}\sigma_{\eta_3}\end{bmatrix} \tag{12}$$

Now, because we are only able to measure the difference between latent utilities (because the choice probabilities depend only on the differences between latent utilities), these parameters are underidentified (because they define the latent utilities for all the brands). We thus have to impose 6 restrictions on the above parameters. In order to make comparisons with the previous results we chose to impose the following restrictions: $\rho_{\varepsilon^b_{13}} = 0$, $\rho_{\varepsilon^b_{23}} = 0$, $\sigma_{\varepsilon^b_3} = \sigma_{\varepsilon^b_1}$, $\rho_{\varepsilon^b_3\eta_1} = 0$, $\rho_{\varepsilon^b_3\eta_2} = 0$, and $\rho_{\varepsilon^b_3\eta_3} = \rho_{\varepsilon^b_1\eta_3}$.

The log-likelihood function is now changed to

Table 3
Greater Number of Draws

	Complete Model	$\rho = 0$	$\rho = 0, \sigma_{\varepsilon^b} = 0$
Yogurt Market			
Dannon	2.65	1.66	1.46
	(0.29)	(0.13)	(0.08)
Yoplait	2.77	1.76	1.52
	(0.31)	(0.15)	(0.10)
Price	-10.45	-7.06	-6.31
	(1.00)	(0.47)	(0.31)
Display	0.14	0.17	0.23
	(0.08)	(0.08)	(0.04)
Feature	0.004	-0.058	-0.018
	(0.04)	(0.04)	(0.03)
Loyalty	0.79	0.83	0.83
	(0.04)	(0.04)	(0.04)
σ_η	0.039	0.038	0.038
	(0.002)	(0.002)	(0.002)
σ_{ε^b}	0.29	0.19	
	(0.04)	(0.02)	
ρ	0.79		
	(0.06)		
$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}$	0.69	0.71	0.68
	(0.03)	(0.03)	(0.03)
$\sigma_{\varepsilon^a_{3-1}}$	0.83	0.88	0.88
	(0.04)	(0.05)	(0.04)
LL	-1079.3	-1095.6	-1126.6
Ketchup Market			
DelMonte	-1.46	-1.34	-1.11
	(0.16)	(0.15)	(0.11)
Hunts	-1.45	-1.35	-1.16
	(0.15)	(0.13)	(0.10)
Price	-2.62	-2.37	-2.20
	(0.32)	(0.30)	(0.24)
Display	0.57	0.58	0.49
	(0.06)	(0.06)	(0.04)
Feature	0.39	0.34	0.42
	(0.09)	(0.09)	(0.05)
Loyalty	0.57	0.57	0.61
	(0.05)	(0.05)	(0.04)
σ_η	0.056	0.056	0.056
	(0.001)	(0.001)	(0.001)
σ_{ε^b}	0.26	0.25	
	(0.03)	(0.03)	
ρ	0.28		
	(0.14)		
$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}$	0.64	0.60	0.46
	(0.09)	(0.10)	(0.10)
$\sigma_{\varepsilon^a_{3-1}}$	0.45	0.44	0.57
	(0.05)	(0.05)	(0.05)
LL	123.9	121.2	96.4

LL is the value of the log likelihood minus the fixed terms. The number of draws is 20 for the GHK and 100 for ε^b_{jt}.

$$L(\beta, \alpha_0, \alpha_1, V_{\varepsilon^b}, V_\eta, V_{\varepsilon^b \eta}) = \prod_{t=1}^{T} f(\eta_{ht} \forall h; V_\eta; \alpha_0, \alpha_1) g(d_{it} \; \forall i \in I_t \mid \eta_{jt} \; \forall j) \tag{13}$$

where $f(\eta_{ht} \forall h; V_\eta; \alpha_0, \alpha_1)$ is the density function of a normal random vector with mean zero and variance V_η, and $g(d_{it} \; \forall i \in I_t \mid \eta_{jt} \; \forall j)$ is the probability of observing choices $d_{it} \; \forall i \in I_t$ in week t given $\eta_{jt} \; \forall j \in J$. Finally, $g(d_{it} \; \forall i \in I_t \mid \eta_{jt} \; \forall j)$ can be written as

$$g(d_{it} \; \forall i \in I_t \mid \eta_{jt} \; \forall j) = \int \int \cdots \int \prod_{i \in I_t} Prob(j_{it}; i, t, \varepsilon^b_{jt} \forall j, t) f(\varepsilon^b_{kt} \forall k \mid \eta_{kt} \forall k; V_{\varepsilon^b \eta}, V_{\varepsilon^b}, V_\eta) \prod_{k=1}^{J} d\varepsilon^b_{kt} \tag{14}$$

where $f(\varepsilon^b_{kt} \forall k \mid \eta_{kt} \forall k; V_{\varepsilon^b \eta}, V_{\varepsilon^b}, V_\eta)$ is the density of the vector ε^b_t given η_t.

Testing for endogeneity is testing for $V_{\varepsilon^b \eta} = 0$. Testing for common demand shocks across consumers is testing for $V_{\varepsilon^b} = 0$. The results are presented in Tables 4 and 5.

These results basically confirm the results obtained above: common errors in the choice model and endogeneity in the marketing-mix variables (i.e., price) matter substantially. Furthermore, these results show that the constraints that were relaxed in this generalization (allowing for richer V_η, V_{ε^b}, and $V_{\varepsilon^b \eta}$) are rejected, i.e., it is important to allow for rich error structures. However, the parameters of the choice model do not change substantially. Given this fact and in the interest of easier computability, we restrict attention in the following sections to the constrained probit model (constraints on V_η, V_{ε^b}, and $V_{\varepsilon^b \eta}$).[16]

4. Endogeneity in the Slope

Another potential and serious problem is endogeneity resulting from the parameters of the marketing-mix variables (endogeneity in the slope). If the firms realize the existence of an increased consumer sensitivity to price, it may be a good strategy to cut prices and this will ultimately be observed in the data. We restrict our attention to endogeneity in the slope in order to simplify the computations.[17] The latent utility model is assumed to be

$$U_{ijt} \; = \; X'_{-P,ijt} \beta_{-P} \; + \; P_{jt}(\beta_P + \varepsilon^d_{jt}) \; + \; \varepsilon^a_{ijt} \tag{15}$$

where X_{-P} represents all the explanatory variables X except for the variable price, β_{-P} represents the vector β except for the parameter associated with the variable price which is represented by β_P. The pricing decision rule is represented by equation (2). All error terms are independent except

Table 4
Generalized Model

	Complete Model	Zero Correlations	Zero Common Errors
Yogurt Market			
Dannon	2.35	1.62	1.49
	(0.25)	(0.13)	(0.09)
Yoplait	2.48	1.73	1.55
	(0.26)	(0.14)	(0.10)
Price	-9.46	-6.92	-6.40
	(0.88)	(0.47)	(0.35)
Display	0.21	0.22	0.24
	(0.09)	(0.08)	(0.05)
Feature	0.018	-0.030	-0.016
	(0.05)	(0.04)	(0.03)
Loyalty	0.86	0.84	0.85
	(0.05)	(0.04)	(0.04)
σ_{η_1}	0.053	0.053	0.053
	(0.005)	(0.006)	(0.006)
$\rho_{\eta_{12}}$	0.0019	0.0078	0.0078
	(0.23)	(0.20)	(0.20)
$\rho_{\eta_{13}}$	0.072	0.030	0.030
	(0.14)	(0.14)	(0.14)
σ_{η_2}	0.024	0.024	0.024
	(0.001)	(0.001)	(0.001)
$\rho_{\eta_{23}}$	0.056	0.040	0.040
	(0.16)	(0.16)	(0.16)
σ_{η_3}	0.031	0.031	0.031
	(0.003)	(0.002)	(0.002)
$\sigma_{\varepsilon_1^b}$	0.20	0.14	
	(0.04)	(0.03)	
$\rho_{\varepsilon_{12}^b}$	-0.20	0.18	
	(0.29)	(0.15)	
$\sigma_{\varepsilon_2^b}$	0.26	0.22	
	(0.05)	(0.05)	
$\rho_{\varepsilon_1^b \eta_1}$	0.66		
	(0.17)		
$\rho_{\varepsilon_1^b \eta_2}$	0.030		
	(0.39)		
$\rho_{\varepsilon_1^b \eta_3}$	0.24		
	(0.28)		
$\rho_{\varepsilon_2^b \eta_1}$	-0.36		
	(0.16)		
$\rho_{\varepsilon_2^b \eta_2}$	0.62		
	(0.19)		
$\rho_{\varepsilon_2^b \eta_3}$	-0.015		
	(0.21)		
$\rho_{\varepsilon_{2-1}^a \varepsilon_{3-1}^a}$	0.70	0.71	0.67
	(0.04)	(0.03)	(0.03)
$\sigma_{\varepsilon_{3-1}^a}$	0.93	0.91	0.90
	(0.06)	(0.05)	(0.05)
LL	-1039.6	-1061.0	-1089.0

LL is the value of the log likelihood minus the fixed terms. The number of draws is 10 for the GHK and 10 for ε_{jt}^b.

Table 5
Generalized Model

	Complete Model	Zero Correlations	Zero Common Errors
Ketchup Market			
Dannon	-1.31	-1.14	-1.10
	(0.17)	(0.14)	(0.11)
Yoplait	-1.32	-1.16	-1.16
	(0.16)	(0.13)	(0.11)
Price	-2.50	-2.17	-2.19
	(0.38)	(0.30)	(0.25)
Display	0.43	0.47	0.49
	(0.07)	(0.05)	(0.04)
Feature	0.47	0.35	0.42
	(0.11)	(0.08)	(0.05)
Loyalty	0.56	0.57	0.60
	(0.04)	(0.05)	(0.05)
$\sigma_{\eta 1}$	0.029	0.029	0.029
	(0.002)	(0.001)	(0.001)
$\rho_{\eta 12}$	0.038	0.019	0.019
	(0.24)	(0.24)	(0.24)
$\rho_{\eta 13}$	-0.027	-0.062	-0.062
	(0.13)	(0.11)	(0.11)
$\sigma_{\eta 2}$	0.081	0.081	0.081
	(0.004)	(0.005)	(0.005)
$\rho_{\eta 23}$	0.025	0.016	0.016
	(0.18)	(0.19)	(0.19)
$\sigma_{\eta 3}$	0.044	0.044	0.044
	(0.002)	(0.002)	(0.002)
$\sigma_{\varepsilon^b_1}$	0.20	0.19	
	(0.03)	(0.03)	
$\rho_{\varepsilon^b_{12}}$	-0.24	-0.97	
	(0.17)	(0.14)	
$\sigma_{\varepsilon^b_2}$	0.30	0.21	
	(0.10)	(0.08)	
$\rho_{\varepsilon^b_1 \eta 1}$	0.64		
	(0.10)		
$\rho_{\varepsilon^b_1 \eta 2}$	0.38		
	(0.17)		
$\rho_{\varepsilon^b_1 \eta 3}$	0.54		
	(0.17)		
$\rho_{\varepsilon^b_2 \eta 1}$	0.44		
	(0.21)		
$\rho_{\varepsilon^b_2 \eta 2}$	0.51		
	(0.19)		
$\rho_{\varepsilon^b_2 \eta 3}$	0.40		
	(0.22)		
$\rho_{\varepsilon^a_{2-1} \varepsilon^a_{3-1}}$	0.62	0.53	0.45
	(0.10)	(0.11)	(0.11)
$\sigma_{\varepsilon^a_{3-1}}$	0.47	0.49	0.57
	(0.05)	(0.05)	(0.05)
LL	188.9	178.6	161.6

LL is the value of the log likelihood minus the fixed terms. The number of draws is 10 for the GHK and 10 for ε^b_{jt}.

for the pairs (ε^d_{jt},η_{jt}) $\forall j$ for which there is a correlation ρ_d. Given the interpretation above for the possible endogeneity in the slope, we expect ρ_d to be positive.

The likelihood function can then be constructed similarly to the one in the previous section (the base case) taking now into account the fact that the integration in function g(.) is performed over the variables ε^d_{jt}, and ρ is substituted for ρ_d.

Testing for endogeneity is testing for $\rho_d = 0$. The results for the yogurt and the ketchup markets are presented in Table 6.

These results confirm the results obtained in the base case and show their robustness to allowing for endogeneity in the slope. They show clearly the importance of considering the endogeneity of the pricing decision: (i) there are important common shocks across individuals (the hypothesis $\sigma_{\varepsilon^d} = 0$ is rejected), (ii) there is significant endogeneity of the price variable (the hypothesis $\rho_d = 0$ is rejected), and (iii) the parameter that estimates the impact of price on choice seems to be clearly underestimated if endogeneity is not taken into account (and as expected, the underestimation is larger in this case). When accounting for endogeneity in the slope, the price effect is underestimated by 87% for the yogurt market; for the ketchup market, it is underestimated by 22%. However, these results are merely exploratory and it would be important to allow for both types of common shocks (additive and in the slope) and allow the data to indicate which type is more important.

5. Heterogeneity in the Consumer Pool

Another important issue that has been widely considered in the literature on brand choice is the existence of unobserved heterogeneity in the consumer pool (for example, Gonul and Srinivasan 1993). This heterogeneity was not considered in the previous sections. It is possible that the previous models accounting for endogeneity were simply picking up this unobserved heterogeneity. As a result, we include heterogeneity in the slope and then test whether accounting for endogeneity is still important.

The latent utility model is assumed to be

$$U_{ijt} \;=\; X'_{-P,ijt}\beta_{-P} \;+\; P_{jt}(\beta_P + \varepsilon^h_i) \;+\; \varepsilon^a_{ijt} \;+\; \varepsilon^b_{jt} \tag{16}$$

where ε^h_i represents the household specific shock in the price effect, β_P represents the average price effect across households, and $\beta_P + \varepsilon^h_i$ represents the price effect for household i. The pricing decision rule is again represented by equation (2). The disturbance ε^h_i is independent of all the other errors in the model and is normally distributed with mean zero and variance $\sigma^2_{\varepsilon^h}$.

Table 6
Endogeneity in the Slope

	Complete Model	$\rho_d = 0$	$\rho_d = 0, \sigma_{\varepsilon^d} = 0$
Yogurt Market			
Dannon	3.09	1.65	1.49
	(0.34)	(0.11)	(0.08)
Yoplait	3.19	1.74	1.56
	(0.36)	(0.12)	(0.09)
Price	-11.99	-7.03	-6.41
	(1.19)	(0.41)	(0.30)
Display	0.19	0.23	0.24
	(0.07)	(0.07)	(0.05)
Feature	0.039	-0.026	-0.016
	(0.04)	(0.04)	(0.03)
Loyalty	0.89	0.86	0.85
	(0.04)	(0.04)	(0.04)
σ_η	0.039	0.038	0.038
	(0.002)	(0.002)	(0.002)
σ_{ε^d}	0.58	0.30	
	(0.09)	(0.05)	
ρ_d	0.90		
	(0.03)		
$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}$	0.66	0.70	0.67
	(0.04)	(0.03)	(0.03)
$\sigma_{\varepsilon^a_{3-1}}$	0.93	0.92	0.90
	(0.05)	(0.05)	(0.04)
LL	-1091.1	-1112.9	-1133.6
Ketchup Market			
DelMonte	-1.41	-1.28	-1.10
	(0.15)	(0.14)	(0.11)
Hunts	-1.41	-1.30	-1.16
	(0.14)	(0.13)	(0.10)
Price	-2.66	-2.39	-2.19
	(0.32)	(0.30)	(0.24)
Display	0.59	0.60	0.50
	(0.05)	(0.05)	(0.04)
Feature	0.42	0.37	0.42
	(0.08)	(0.07)	(0.05)
Loyalty	0.61	0.62	0.60
	(0.05)	(0.05)	(0.04)
σ_η	0.056	0.056	0.056
	(0.001)	(0.001)	(0.001)
σ_{ε^d}	0.23	0.22	
	(0.03)	(0.03)	
ρ_d	0.29		
	(0.11)		
$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}$	0.59	0.55	0.45
	(0.10)	(0.10)	(0.10)
$\sigma_{\varepsilon^a_{3-1}}$	0.53	0.53	0.57
	(0.05)	(0.05)	(0.05)
LL	112.5	109.7	94.9

LL is the value of the log likelihood minus the fixed terms. The number of draws is 10 for the GHK and 10 for ε^d_{jt}.

The log-likelihood function is changed from the one in Section 3 and is now

$$L(\beta, \alpha_0, \alpha_1, \sigma^2_{\varepsilon^b}, \sigma^2_\eta, \rho, \sigma^2_{\varepsilon^h}) = \int\int \cdots \int \prod_{t=1}^{T} \prod_{h=1}^{J} f(\eta_{ht}; \sigma^2_\eta; \alpha_0, \alpha_1) g(d_{it} \; \forall i \in I_t \mid \eta_{jt} \; \forall j) \times \prod_{i=1}^{H} \frac{\phi(\frac{\varepsilon_i^h}{\sigma_{\varepsilon^h}})}{\sigma_{\varepsilon^h}} \prod_{i=1}^{H} d\varepsilon_i^h \quad (17)$$

where H is the number of households in the sample, and $\phi(.)$ is the density function of the standardized normal distribution.

The heterogeneity integrals were also simulated by taking an average across draws of ε_i^h. The results are presented in the Table 7 for the yogurt ($H = 562$) and the ketchup ($H = 252$) markets.

The results show that even accounting for heterogeneity, endogeneity is significant since the null hypothesis $\rho = 0$ is rejected for both categories. Not accounting for endogeneity still results in the price effect being underestimated. Note also that columns 1 through 3 of Table 2 present the results of the model of Table 7 with the restriction $\sigma_{\varepsilon^h} = 0$. Then, by comparing the LL value in the third column of Table 2 with the LL value in the third column of Table 7, one can conclude that when both endogeneity and the common shocks are not in the model, the effects of heterogeneity are statistically significant for the ketchup market. However, when endogeneity is accounted for, the heterogeneity effects disappear (compare the LL value in the first column of Table 2 with the LL value in the first column of Table 7). This might indicate that heterogeneity, when included alone in the model (i.e., without accounting for endogeneity), may pick up some of the endogeneity effects. However, these results have to be interpreted with care because of the limited number of draws in the analysis.

6. Other Pricing Decision Rules

In order to check the robustness of the results to other formulations of the pricing decision rule (i.e., instruments for price), we allowed equation (2), the pricing equation, to include other explanatory variables (i.e., instruments for price). The choice model is the probit one whose results are shown in Table 2.

One pricing rule that has been used in the past in several applications, for example, Winer (1986), is to have lagged market share and lagged price variables in the pricing equation. Equation (2) is then changed to

Table 7
Heterogeneity

	Complete Model	$\rho = 0$	$\rho = 0, \sigma_{\varepsilon^b} = 0$
Yogurt Market			
Dannon	3.11	1.56	1.48
	(0.30)	(0.09)	(0.07)
Yoplait	3.22	1.64	1.55
	(0.32)	(0.11)	(0.09)
Price	-11.98	-6.66	-6.38
	(0.98)	(0.36)	(0.29)
Display	0.15	0.19	0.23
	(0.06)	(0.06)	(0.05)
Feature	0.050	-0.031	-0.011
	(0.03)	(0.04)	(0.04)
Loyalty	0.78	0.84	0.84
	(0.04)	(0.04)	(0.04)
σ_η	0.041	0.038	0.038
	(0.002)	(0.002)	(0.002)
σ_{ε^b}	0.33	0.14	
	(0.05)	(0.02)	
ρ	0.95		
	(0.02)		
$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}$	0.66	0.69	0.67
	(0.04)	(0.03)	(0.03)
$\sigma_{\varepsilon^a_{3-1}}$	0.81	0.90	0.90
	(0.04)	(0.05)	(0.04)
σ_{ε^h}	0.026	0.029	0.026
	(0.24)	(0.26)	(0.28)
LL	-1093.7	-1110.2	-1133.2
Ketchup Market			
DelMonte	-1.46	-1.31	-1.11
	(0.17)	(0.14)	(0.10)
Hunts	-1.48	-1.34	-1.17
	(0.16)	(0.13)	(0.10)
Price	-2.99	-2.68	-2.19
	(0.36)	(0.31)	(0.24)
Display	0.45	0.47	0.49
	(0.05)	(0.05)	(0.04)
Feature	0.45	0.39	0.43
	(0.07)	(0.06)	(0.05)
Loyalty	0.58	0.58	0.60
	(0.04)	(0.04)	(0.04)
σ_η	0.056	0.056	0.056
	(0.001)	(0.001)	(0.001)
σ_{ε^b}	0.25	0.22	
	(0.03)	(0.03)	
ρ	0.33		
	(0.16)		
$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}$	0.48	0.41	0.49
	(0.14)	(0.15)	(0.09)
$\sigma_{\varepsilon^a_{3-1}}$	0.49	0.48	0.56
	(0.05)	(0.05)	(0.04)
σ_{ε^h}	0.080	0.081	0.081
	(0.06)	(0.06)	(0.06)
LL	112.0	109.5	98.6

LL is the value of the log likelihood minus the fixed terms. The number of draws is 10 for the GHK, 10 for ε^b_{jt}, and 10 for ε^h_i.

$$P_{jt} = \alpha_{j0} + \alpha_{j1} P_{j,t-1} + \alpha_{j2} S_{j,t-1} + \eta_{jt} \tag{18}$$

where S_{jt} is the market share for brand j in period t, and α_{j2} is a parameter. The rationale behind this model is that managers adjust their prices based on their market performance in the prior period.

The results, presented in Table 8, show that the main conclusions from the previous sections are robust to including lagged market share as an explanatory variable: the common shocks across consumers are statistically significant and there is endogeneity in the price variable. The price parameter in Table 8 is also very close to the one in Table 2 (which presents the results with the simplified pricing rule).

Another important alternative to consider was to take lagged prices out of the pricing rule and use only other variables as instruments for current prices. We had available some cost variables for both markets that could serve as instruments. For yogurt, the cost variables were prices of milk (used to make yogurt) in different regions of the country. For ketchup the cost variables were prices of different types of tomatoes.[18] The cost variables had statistically significant explanatory power on current prices. The results, available upon request from the authors, strengthened, as expected, the main message of the results presented above: not accounting for endogeneity may result in a substantial underestimation of the price coefficient. The strengthening of the results is expected because lagged prices are possibly still correlated with the current utility equation residuals.

7. Conclusion

Our results show the importance of accounting for endogeneity in brand choice models estimated using scanner panel data from two product categories. Variables which are unobserved to the marketing researchers seem to be used by marketing managers to set the levels of the decision variables and to affect consumer brand choice. The results in this paper show that not accounting for endogeneity may result in underestimation of the market response to price. We also showed that these results are robust to more complex error structures, to a formulation incorporating endogeneity of the slope, to heterogeneity in the consumer pool, and to relaxations in the pricing rule.

There are several limitations to this study. First, we tested for endogeneity in only two product categories with a limited number of alternatives. Tests for endogeneity in other product categories allowing for a greater number of alternatives are still needed. Second, the models of the firms' decision making are relatively ad-hoc. In future work, one may want to start with a theory about firms' behavior and interactions and then derive the firms' decision making with more clear restrictions

Table 8
Lagged Market Share in Pricing Rule

	Complete Model	$\rho = 0$	$\rho = 0, \sigma_{\varepsilon^b} = 0$
Yogurt Market			
Dannon	2.61	1.60	1.49
	(0.25)	(0.11)	(0.08)
Yoplait	2.70	1.70	1.56
	(0.27)	(0.11)	(0.10)
Price	-10.27	-6.81	-6.41
	(0.89)	(0.39)	(0.31)
Display	0.22	0.23	0.24
	(0.08)	(0.08)	(0.05)
Feature	0.019	-0.033	-0.016
	(0.04)	(0.04)	(0.03)
Loyalty	0.80	0.83	0.85
	(0.04)	(0.04)	(0.04)
σ_η	0.038	0.037	0.037
	(0.002)	(0.002)	(0.002)
σ_{ε^b}	0.28	0.16	
	(0.04)	(0.02)	
ρ	0.82		
	(0.05)		
$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}$	0.68	0.71	0.67
	(0.04)	(0.03)	(0.03)
$\sigma_{\varepsilon^a_{3-1}}$	0.83	0.90	0.90
	(0.04)	(0.05)	(0.05)
LL	-1085.4	-1103.1	-1126.3
Ketchup Market			
DelMonte	-1.32	-1.18	-1.10
	(0.14)	(0.13)	(0.11)
Hunts	-1.31	-1.20	-1.16
	(0.13)	(0.12)	(0.10)
Price	-2.53	-2.25	-2.19
	(0.29)	(0.28)	(0.24)
Display	0.52	0.53	0.50
	(0.04)	(0.04)	(0.04)
Feature	0.41	0.36	0.42
	(0.08)	(0.07)	(0.05)
Loyalty	0.58	0.59	0.60
	(0.05)	(0.05)	(0.04)
σ_η	0.055	0.055	0.055
	(0.001)	(0.001)	(0.001)
σ_{ε^b}	0.24	0.22	
	(0.03)	(0.03)	
ρ	0.32		
	(0.12)		
$\rho_{\varepsilon^a_{2-1}\varepsilon^a_{3-1}}$	0.58	0.53	0.49
	(0.10)	(0.10)	(0.09)
$\sigma_{\varepsilon^a_{3-1}}$	0.50	0.51	0.56
	(0.05)	(0.05)	(0.05)
LL	115.2	111.5	99.8

LL is the value of the log likelihood minus the fixed terms. The number of draws is 10 for the GHK, and 10 for ε^b_{jt}.

between the supply and the demand side of the model. Third, we only analyzed endogeneity in the price variable. Several other explanatory variables are also probably affected by endogeneity, for example, displays and features. Fourth, several aspects of the execution of our tests could be improved upon with greater computer capabilities: greater number of simulations, allowing for both additive and slope endogeneity at the same time, allowing for more general heterogeneity structures, including in the model both very general error structures, endogeneity in the slope, and heterogeneity. Fifth, one could check the robustness of the results to other estimation methods, in particular, to estimation methods that rely on less distribution assumptions. Finally, recent developments in the literature on brand choice such as the when/what/how much decision (Gupta 1988, Chiang 1991, Chintagunta 1993) could also be included in the model.[19]

Nevertheless, we believe that our results may be robust to all these extensions of the base model: endogeneity in the marketing-mix variables may be important and can create a substantial bias in the parameter estimates in several markets. Researchers in the brand choice area should now question very seriously the existence of endogeneity in their models, and if not accounting for it, they should at least test for its existence.[20]

REFERENCES

ALLENBY, G. AND P. ROSSI (1991), "There is No Aggregation Bias: Why Macro Logit Models Work," *Journal of Business and Economic Statistics,* **9**, 1-14.

BASS, F. (1969), "A Simultaneous Equation Regression Study of Advertising and Sales of Cigarettes," *Journal of Marketing Research,* **6**, 291-300.

BERRY, S. (1994), "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics,* **25**, 242-262.

BERRY, S., J. LEVINSOHN AND A. PAKES (1995), "Automobile Prices in Market Equilibrium," *Econometrica,* **63**, 841-890.

BERRY, S., M. CARNALL AND P. SPILLER (1995), "Airline Hubs: Costs, Markups and the Implications of Customer Heterogeneity," *mimeo,* University of California at Berkeley.

CHIANG, J. (1991), "A Simultaneous Approach to the Whether, What, and How Much to Buy Questions," *Marketing Science,* **10**, 297-315.

CHINTAGUNTA, P. (1993), "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households," *Marketing Science,* **12**, 184-208.

GOLDBERG, P. (1995), "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," *Econometrica,* **63**, 891-952.

GONUL, F. AND K. SRINIVASAN (1993), "Modeling Multiple Sources of Heterogeneity in Multinomial Logit Models: Methodological and Managerial Issues," *Marketing Science,* **12**, 213-229.

GOURIEROUX, C. AND A. MONFORT (1993), "Simulation-Based Inference: a Survey with Special Reference to Panel Data Models," *Journal of Econometrics,* **59**, 5-33.

GUADAGNI, P. AND J. LITTLE (1983), "A Logit Model of brand Choice Calibrated on Scanner Data," *Marketing Science,* **2**, 203-238.

GUPTA, S. (1988), "Impact of Sales Promotion on When, What, and How Much to Buy," *Journal of Marketing Research,* **25**, 342-355.

HAJIVASSILIOU, V., D. MCFADDEN AND P. RUUD (1992), "Simulation of Multivariate Normal Orthant Probabilities: Theoretical and Computational Results," *mimeo,* University of California at Berkeley.

HAJIVASSILIOU, V. AND D. MCFADDEN (1993), "The Method of Simulated Scores for the Estimation of LDV Models with an Application to External Debt Crisis," *mimeo,* University of California, Berkeley.

HECKMAN, J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica,* **46**, 931-959.

HECKMAN, J. (1979), "Sample Selection Bias as a Specification Error," *Econometrica,* **47**, 153-161.

KEANE, M. (1993), *Simulation Estimation for Panel Data Models with Limited Dependent Variables,* in: C. R. Rao, G. S. Maddala and H. D. Vinod, eds., The Handbook of Statistics 11: Econometrics. Amsterdam: Elsevier Science Publishers, pp. 545-572.

MCFADDEN, D. (1973), *Conditional Choice Analysis of Qualitative Choice Behavior,* in Frontiers of Econometrics, P.Zarembka, ed. New York: Academic Press, 105-42.

MILGROM, P. AND J. ROBERTS (1986), "Prices and Advertising Signals of Product Quality," *Journal of Political Economy,* **94**, 796-821.

NEWEY, W. (1985), "Maximum Likelihood Estimation Testing and Conditional Moment Tests," *Econometrica,* **53**, 1047-1070.

NEWEY, W. (1987), "Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables," *Journal of Econometrics,* **36**, 230-251.

NEWEY, W. AND D. MCFADDEN (1994), *Large Sample Estimation and Hypothesis Testing,* in: Z. Griliches and M. Intriligator, eds., Handbook of Econometrics 4. Amsterdam: Elsevier Science Publishers.

RIVERS, D. AND Q. H. VUONG (1988), "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models," *Journal of Econometrics,* **39**, 347-366.

WINER, R. (1986), "A Reference Price Model of Brand Choice for Frequently Purchased Products," *Journal of Consumer Research,* **13**, 250-256.

NOTES

[1]A second source of endogeneity in scanner data applications, not considered here, is due to the way the price variable has been constructed to include coupon redemption. Coupon redemption is also likely to be strongly correlated with the error term (in particular, with the heterogeneity disturbance in any formulation accounting for consumer heterogeneity).

[2]It is still possible that these models can be used for prediction purposes if the managers keep making decisions in the same way as in the data used for estimation.

[3]In those weeks the consumer has utility equal to minus infinity if she does not buy any product, and the utility does not improve if she buys more than one product. We have also performed runs of the model with the no-purchase option: The results are similar to the ones presented her, and our main message is substantially strengthened.

[4]We have also performed runs of the model with the loyalty variable specified as an exponential smoothing of past purchases, as in Guadagni and Little 1983. The results are very similar to the ones presented here.

[5]This is for easier computability. In the following sections we consider ε^a_{ijt} to be normally distributed.

[6]The formulation for the common demand shocks is a random effects one because of the small number of purchases per week. An alternative is a fixed effects formulation.

[7]Note that the instruments being used have to be independent of the residuals in the utility equation. Because of forward buying and stock piling, one might argue that lagged prices could be correlated with the utility equation residuals. However, these forward buying and stock piling effects have been found to be of very limited size (see Gupta 1988). Another potential, and more serious problem, is that the common shock in the utility equation, ε^b_{jt}, may be correlated through time, and then lagged prices would be correlated with the current period common shock. But, if this is the case, the results obtained here would be conservative, and would only underestimate the endogeneity effects. A main advantage of using lagged prices as instruments is that they are readily available to the researcher. Section 6 reports results from the use of other instruments for price.

[8]For the profit maximization interpretation of the pricing rule, the assumption that η_{jt} is independent across t may also not be empirically satisfied. This assumption is kept throughout the paper. Relaxing it would create the type of estimation problems faced in Section 5, the log-likelihood function would not be additively separable across periods.

[9]There were originally 10 stores in this data set. Including more stores in the analysis would only increase the choice set into products where there was a small degree of substitutability for the consumers being considered. The analysis reported in this paper was limited to one store for both the yogurt and the ketchup markets. The results for more than one store are in every way similar to the results presented here and available upon request from the authors.

[10]We did not include more brands in the analysis for computational reasons.

[11]There were 9 stores in this data set.

[12]The value of the maximum log-likelihood can be positive in these cases because the density of η_{jt}, included in the log-likelihood function, can be greater than one (i.e., the log of the density can be greater than zero), and compensate

for the log of the probability of choice of each brand (which is less than zero).

[13]This method allows also for estimation with any number of brands. Notice also that the standard method of simulated moments does not work in this case because the number of alternatives in that method is too large - J raised to the power of the number of consumers that purchased a product in that week (see Hajivassiliou and McFadden 1993).

[14]For the GHK simulator, mentioned below, these estimators are consistent and asymptotically normal if $K/\sqrt{N} \rightarrow \infty$ as $N \rightarrow \infty$, where K is the number of simulations and N is the number of observations.

[15]Note that this is the standard scaling in Probit models, which is different from the scaling used in the previous Section.

[16]Another potentially important generalization of the error structures, not studied here, is to allow the common shocks ε_{jt}^{b} to be autocorrelated through time. With the current computational capabilities and this likelihood approach, this structure is however relatively complicated to estimate because one has both cross-section and time-series correlation of the residuals – as in Section 5 below.

[17]Alternatively, we could have considered both endogeneity in the slope and additive endogeneity, the one we considered in the previous Sections.

[18]All cost variables were obtained from the United States Department of Agriculture.

[19]Other results obtained by the authors in a model with the no-purchase option also show that the results presented above may even underestimate the endogeneity problems.

[20]One easily applicable test is the test of the existence of common shocks in the utility equation, which might be a strong cue for the existence of endogeneity problems. The test is a lagrange multiplier test on the restriction that the variance of the common shocks is zero. Details are available in Newey and McFadden (1994).