SHA YANG and GREG M. ALLENBY*

A consumer's preference for an offering can be influenced by the preferences of others in many ways, ranging from social identification and inclusion to the benefits of network externalities. In this article, the authors introduce a Bayesian spatial autoregressive discrete-choice model to study the preference interdependence among individual consumers. The autoregressive specification can reflect patterns of heterogeneity in which influence propagates within and across networks. These patterns cannot be modeled with standard random-effect specifications and can be difficult to capture with covariates in a linear model. The authors illustrate their model of interdependent preferences with data on automobile purchases and show that preferences for Japanese-made cars are related to geographically and demographically defined networks.

Modeling Interdependent Consumer Preferences

Preferences and choice behavior are influenced by a consumer's own tastes and the tastes of others. People who identify with a particular group often adopt the preferences of that group, which results in interdependent choices. Examples include the preference for particular brands (e.g., Honda Odyssey) and even entire product categories (e.g., minivans). Interdependence may be driven by social concerns, endorsements from respected people who increase a brand's credibility, or learning the preference of others who may have information not available to the decision maker. Moreover, because people engage in multiple activities with their families, coworkers, neighbors, and friends, interdependent preferences can propagate across and through multiple networks.

Quantitative models of consumer purchase behavior often do not recognize that preferences and choices are interdependent. Economic models of choice typically assume that a consumer's latent utility is a function of brand and attribute preferences, not the preferences of others. Preferences are assumed to vary across consumers in a manner described either by exogenous covariates, such as demographics (e.g., household income), or by independent draws from a mixing distribution in random-effects models (see Allenby and Rossi 1999; Kamakura and Russell 1989). However, if preferences in a market are interdependently determined, their pattern will not be well represented by a simple linear model of exogenous covariates. Failure to include high-order interaction terms in the model to reflect interdependent influences results in a correlated structure of unobserved heterogeneity in which the draws from a random-effects mixing distribution are dependent, not independent.

In this article, we introduce a Bayesian model of interdependent preferences in a consumer choice context. We employ a parsimonious autoregressive structure that captures the endogenous relationship between a consumer's preference and the preference of others in the same network. Our model allows for a complex network associated with multiple explanatory variables. In addition, our model enables us to incorporate explanatory variables in both an exogenous and an endogenous (i.e., interdependent) manner. We employ a Markov chain Monte Carlo method to estimate the model parameters.

In the subsequent section, we review the literature in economics and marketing on interdependent preference, which provides rationale for our model structure. In the section "A Hierarchical Bayes Autoregressive Mixture Model," we lay out the model and estimation procedure, and in "A Simulation Study," we present a numerical simulation to assess the accuracy of the model in recovering the true parameter values. We then set forth an empirical application of the model to study the interdependence in a binary choice decision: whether to buy a foreign brand or a domestic brand of midsize car.

*Sha Yang is Assistant Professor of Marketing, Stern School of Business, New York University (e-mail: shayang@stern.nyu.edu). Greg M. Allenby is Kurtz Chair Professor in Marketing, Fisher College of Business, Ohio State University (e-mail: allenby.1@osu.edu). The authors thank three anonymous *JMR* reviewers for their constructive comments and suggestions. The authors also thank seminar participants at University of California, Berkeley; New York University; Cornell University; Hong Kong University of Science and Technology; National University of Singapore; University of Connecticut; Case Western Reserve University; University of Wisconsin, Madison; and Tohoku University, Sendai, Japan, for their helpful comments.

REVIEW OF INTERDEPENDENT PREFERENCE STUDIES

People do not live in a world of isolation; they interact with one another when forming their opinions, beliefs, and preferences. Interdependent preference (or preference interaction) has been defined as "occurring when an agent's preference ordering over the alternatives in a choice set depends on the actions chosen by other agents" (Manski 2000, p. 120). In studies of interdependence, this effect has also been called "peer influences" (Duncan, Haller, and Portes 1968), "neighborhood effects" (Case 1991), "bandwagon effect" (Leibenstein 1950), and "conformity" (Bernheim 1994).

Interdependent preferences can arise in many ways, including social concerns, reduced transaction costs (i.e., network externalities), and the signaling effect of another's brand ownership on inferred attribute levels. In addition, interdependent preferences may appear to be present in economic demand models in which key explanatory variables are either omitted (e.g., household income) or unobserved (e.g., media exposure). Our review of the literature focuses on the former explanation, and the potential mechanisms by which interdependent preference arises guide our model specification. The following review provides a useful guide for researchers studying the origins of utility.

Duesenberry (1949) and Leibenstein (1950) have conducted pioneering work on interdependent preference. Duesenberry describes several examples of interdependence in consumer consumption behavior. Using data on consumer purchases made in 1935 and 1936, he finds that the percentage of income spent on consumption is highly correlated with the person's rank order in the local income distribution. Leibenstein formally incorporates the phenomenon of "conspicuous consumption" into a theory of consumer demand. Through a conceptual experiment, he shows that the demand curve is more elastic if there is a bandwagon effect than if demand is based only on the functional attributes of the commodity. Since these works were published, a large body of literature that examines the theory and empirical evidence of interdependent preference has emerged.

Theoretical Research on Interdependent Preference

In the theory domain, Hayakawa and Venieris (1977) derive several utility theories and axioms for preference interdependence. Their theories predict that the income effect associated with a price change will become dominant as the budget expenditure is relaxed. Moreover, their research points to the need for a concept of psychological complementarity to capture the role of a reference group in a consumer choice calculus. Although Hayakawa and Venieris's framework is mainly static, several other authors have tried to address the influence of preference interdependence by adding a temporal dimension. For example, Cowan, Coman, and Swann (1997) derive the steady state and dynamic properties of the distribution of consumption when different reference groups are used. Bernheim (1994) suggests a theory of how standards of behavior may evolve in response to changes in the distribution of intrinsic preferences.

Although the focus of studying the interdependent effects in economics is mainly on the effects' impact on demand theory and econometric implications, researchers in marketing have paid more attention to explaining the mechanisms of the interdependent preferences among consumers in the context of reference-group formation and influence. These two areas of research are distinguished. The first examines the reasons that individuals conform to the behavior of a reference group. Researchers have identified three sources of social influence on buyer behavior: internalization, identification, and compliance (Burnkrant and Cousineau 1975; Kelman 1961; Lessig and Park 1978). *Internalization* occurs when people adopt other people's influence because they perceive it as "inherently conductive to the maximization of [their] value" (Burnkrant and Cousineau 1975, p. 207). In other words, people are willing to learn from others because they believe it could help them make a better decision that optimizes their own returns. *Identification* occurs when people adopt from others because the "behavior is associated with a satisfying self-defining relationship" to the other (Burnkrant and Cousineau 1975, p. 207). *Compliance* occurs when "the individual conforms to the expectations of another in order to receive a reward or avoid punishment mediated by that other" (Burnkrant and Cousineau 1975, p. 207).

The second area of research in marketing examines the relative influence of alternative mechanisms on individual consumer behavior. Bearden and Etzel (1982) study how reference groups influence an individual consumer's purchase decisions at both the product and the brand level. Lessig and Park (1982) show that the degree of reference-group influence is dependent on product-related characteristics, such as complexity, conspicuousness, and brand distinction. Childers and Rao (1992) find that reference-group influence varies for products consumed in different occasions (public versus private) and for different reference groups (familial versus peer).

Interdependent preferences can therefore be associated with multiple covariates and can lead to either conformity or individuality in preferences. In the next section, we present a model capable of reflecting these effects within the framework of an economic choice model.

Measuring Interdependence

In the empirical domain, a considerable amount of effort has been devoted to developing econometric models and estimation methods that incorporate interpersonal dependence (Kapteyn 1977; Pollak 1976; van de Stadt, Kapteyn, and van de Geer 1985). Some empirical applications include studying interdependent preference in consumer expenditure allocations (Alessie and Kapteyn 1991; Darrough, Pollak, and Wales 1983; Kapteyn et al. 1997), labor supply (Aronsson, Blomquist, and Sacklén 1999), rice consumption (Case 1991), and elections (Smith and LeSage 2000). The focus of these studies is typically aggregate; the dependent variable reflects average behavior (e.g., consumption) within a particular geographic region (e.g., Zip code). Furthermore, the studies tend to use a single network to model the preference interdependence.

Methodological research in marketing on interdependent preferences recently has been spurred by developments in simulation-based estimations that facilitate flexible models of consumer heterogeneity, including models in which dependence is spatially related. Arora and Allenby (1999) propose a conjoint model in which the importance of product attributes in a group decision-making context can be different from the part-worths of each group member. However, their model assumes that the social group is readily identified, and the model does not account for the possibility of multiple networks. Hofstede, Wedel, and Steenkamp (2002) examine the use of alternative spatial prior distributions to geographically smooth model parameters in a study of retail store attributes. In their analysis, they assume that response coefficients in a geographic area are similar to those in neighboring areas. Bronnenberg and Mahajan (2001) combine an autoregressive spatial prior on market shares with a temporal autoregressive process to study variation in the effectiveness of promotional variables in geographically defined markets. Bronnenberg and Sismeiro (2002) use a spatial model to forecast brand sales in markets for which only limited information exists. None of these studies, except for that by Arora and Allenby, is developed to study preferences at the level of the individual consumer.

In this article, we develop a spatial autoregressive mixture model and apply it to the latent utility in a discrete-choice model. The autoregressive model relates a consumer's latent utility to the utility of other consumers, which reflects the potential interdependence of preferences. We incorporate explanatory variables into the autoregressive process through a weighting matrix that describes the network. The mixture aspect of the model enables the weighting matrix to be defined by multiple covariates, and we estimate covariate importance from the data. Covariates are also related to the expected value of the latent utilities to capture exogenous rather than endogenous effects.

A HIERARCHICAL BAYES AUTOREGRESSIVE MIXTURE MODEL

In this section, we first introduce a binary choice model that captures the potential social dependency of preferences among consumers. We then briefly describe the prior distribution specification and estimation procedure using Markov chain Monte Carlo methods. A detailed description of the estimation algorithm is provided in Appendix A.

An Autoregressive Discrete-Choice Model

Suppose we observe choice information for a set of consumers ($i = 1, \ldots, m$) who are not associated with an interdependent network and whose preferences are exogenously determined. Assume that a consumer is observed to make a selection between two choice alternatives ($y_i = 1$ or 0) that is driven by the difference in latent utilities, U_{ik}, for the two alternatives ($k = 1,2$). The probability of the consumer selecting the second alternative over the first is as follows:

$$\Pr(y_i = 1) = \Pr(U_{i2} > U_{i1}) = \Pr(z_i > 0), \tag{1}$$

$$z_i = x_i'\beta + \varepsilon_i, \tag{2}$$

$$\varepsilon_i \sim \text{normal}(0,1), \tag{3}$$

where z_i is the latent preference for the second alterative over the first alternative, x_i is a vector of covariates that captures the differences of the characteristics between the two choice alternatives and characteristics of the consumer, β is the vector of coefficients associated with x_i, and ε_i reflects unobservable factors modeled as error. For example, preference for a durable offering may be dependent on the existence of local retailers that can provide repair service when needed. This type of exogenous preference dependence is well represented by Equation 2. We assume that the error term is independently distributed across consumers, which reflects the absence of interdependent effects. The scale of the error term is equal to one to identify the model coefficients, β, statistically. Stacking the latent preferences, z_i, into a vector results in a multivariate specification:

$$z \sim \text{normal}(X\beta, I). \tag{4}$$

The presence of interdependent networks creates preferences that are endogenous and mutually dependent, resulting in an error covariance matrix (Σ) with nonzero off-diagonal elements. The presence of off-diagonal elements in the covariance matrix leads to conditional and unconditional expectations of preferences that differ. The expectation of latent preference, z, in Equation 4 is equal to $X\beta$, regardless of whether preferences of other consumers are known. However, if the off-diagonal elements of the covariance matrix are nonzero, the conditional expectation of the latent preference for one consumer is correlated with the revealed preference of another consumer:

$$E[z_2|z_1] = X_2'\beta + \Sigma_{21}\Sigma_{11}^{-1}(z_1 - X_1\beta), \tag{5}$$

where the subscripts apply to appropriate elements of the parameters. Positive covariance leads to a greater expectation of preference z_2 if it is known that z_1 is greater than its mean, $X_1\beta$. In the probit model, choice (y_i) is revealed, corresponding to a range of latent preferences (i.e., $z_i > 0$). The computation of conditional expectation is therefore associated with an integration over a range of possible conditioning arguments.

An approach to inducing covariation among the error terms is to augment the error term in Equation 2 with a second error term from an autoregressive process (LeSage 2000):

$$z_i = x_i'\beta + \varepsilon_i + \theta_i\,, \tag{6}$$

$$\theta = \rho W\theta + u, \tag{7}$$

$$\varepsilon \sim N(0,I), \tag{8}$$

$$u \sim N(0,\sigma^2 I), \tag{9}$$

where ε and u are i.i.d. error terms, and θ is a vector of autoregressive parameters where the matrix ρW reflects the interdependence of preferences across consumers. The specification in Equation 7 is similar to that encountered in time-series analysis, except that codependence can exist between two elements, whereas in time-series analysis the dependence is directional (e.g., an observation at time t-k can affect an observation at time t, but not vice versa). Codependence is captured by nonzero entries that appear in both the upper and the lower triangular submatrices of W. We assume that the diagonal elements w_{ii} are equal to zero and that each row sums to one. The coefficient ρ measures the degree of overall association among the units of analysis beyond that captured by the covariates, X. Positive (negative) value of ρ indicates positive (negative) correlation among consumers.

It is worth noting that in our model of interdependent preferences, there is a network propagation effect captured in Equation 7, whereas in the exogenous model the effect associated with a covariate does not propagate among consumers. Our model presents a simple test on the existence of propagation effect. If ρ is significantly different from zero, we conclude that there could be some interdependent preference beyond what is captured in the $x'\beta$ term in Equation 6.

The augmented-error model results in latent preferences with nonzero covariance:

$$z \sim \text{normal}[X\beta, I + \sigma^2(I - \rho W)^{-1}(I - \rho W')^{-1}]. \tag{10}$$

This specification is different from that encountered in standard spatial data models (see Cressie 1991, p. 441), in which the error term ε is not present and the covariance term is equal to $\sigma^2(I - \rho W)^{-1}(I - \rho W')^{-1}$.[1] The advantage of specifying the error in two parts is that it leads itself to estimation and analysis using the method of data augmentation (Tanner and Wong 1987). The parameter θ is responsible for the nonzero covariances in the latent preferences, z, but is not present in the likelihood specification (Equation 10). By augmenting the parameters space with θ, we isolated the effects of the nonzero covariances and simplified the evaluation of the likelihood function.

The elements of the autoregressive matrix, $W = \{w_{ij}\}$, reflect the potential dependence between units of analysis. A critical part of the autoregressive specification pertains to the construction of W. For example, spatial models have employed a coding scheme by which the unnormalized elements of the autoregressive matrix equal one if i and j are neighbors and equal zero otherwise (see Bronnenberg and Mahajan 2001). An alternative specification for the model could involve other metrics, such as Euclidean and Manhattan distances. However, as we noted previously, interdependent preferences can be determined by multiple networks. It is therefore important to allow for a specification of the autoregressive matrix W with multiple covariates.

We specify the autoregressive matrix W as a finite mixture of coefficient matrices, each related to a specific covariate:

$$W = \sum_{k=1}^{K} \phi_k W_k, \tag{11}$$

$$\sum_{k=1}^{K} \phi_k = 1, \tag{12}$$

where k indexes the covariates, k = 1, …, K. The weights, ϕ_k, reflect the relative importance of the component matrices, W_k, and each is associated with a different explanatory variable. For example, W_1 may be related to the physical proximity of individual consumer residences, W_2 to age, W_3 to income, W_4 to ethnicity, and so on. Within each matrix W_k, we assume that the diagonal elements are equal to zero, the off-diagonal elements reflect the distance between consumers in terms of the kth covariate, and each row sums to one. The weighted sum of the component matrices, W, also has these properties because the weights, ϕ_k, sum to one. In our specification, we reparameterize ϕ_k with a logit specification,

$$\phi_j = \frac{\exp(\alpha_j)}{\sum_{k=1}^{K} \exp(\alpha_k)}, \tag{13}$$

and estimate α_j unrestricted, with $\alpha_K = 0$.

The model Equations 1 and 6–9 describe is statistically identified. This is evident when considering the choice probability $\Pr(y_i = 1) = \Pr(z_i > 0) = \Pr(x_i'\beta + \varepsilon_i + \theta_i > 0)$. The right-hand side of the latter expression is zero, and the variance of ε_i is one. These specifications identify the probit model in terms of location and scale, because an arbitrary constant cannot be added to the right-hand side of the expression and multiplying by a scalar quantity would alter the variance of ε_i. Moreover, the finite mixture specification Equations 10–12 describe is identified because the rows of W_k and the mixture probabilities are normalized to add to one. However, we note that care must be exercised in comparing estimates of β from the independent probit model (Equation 4) with those from the autoregressive model (Equation 10) because of the differences in the magnitude of the covariance matrix. We return to the issue of statistical identification in the subsequent section, where we demonstrate properties of the model in a simulation study.

Our model is based on the framework Smith and LeSage (2000) develop, but it extends theirs in the following ways: First, the matrix W in their model is constrained to be only geographically specific. We introduce a more flexible structure to capture different sources of interdependence across multiple networks. Second, most applications in this area of research, including Smith and LeSage (2000), focus on Zip code–level analysis. Our analysis investigates interdependence at the individual consumer level, which is of interest in marketing.

Prior Specification and Markov Chain Estimation

We estimate the autoregressive discrete-choice model using Markov chain Monte Carlo methods. This method of estimation requires specification of prior distributions for the model parameters in Equation 10 and derivation of the full conditional distribution of model parameters. We set the prior distributions to be diffuse and conjugate when possible. We use some standard prior distribution specifications as follows:

$$\beta \sim N(\beta_0, V_\beta), \tag{14}$$

$$\rho \sim U\left[\frac{1}{\lambda_{min}}, \frac{1}{\lambda_{max}}\right], \tag{15}$$

$$\alpha \sim N(\alpha_0, V_\alpha), \tag{16}$$

$$\sigma^2 \sim IG(s_0, q_0). \tag{17}$$

Here, α and β have normal conjugate prior distributions with means set to zero and covariance matrices set to 100I, where I is the identity matrix and σ^2 is assigned a conjugate inverted gamma prior, where $s_0 = 5$ and $q_0 = .1$. We employ

[1]With this model specification, we assume that θ has smaller variance for consumers in larger networks. We believe this property is justified. As the size of the network increases, interdependence implies that there exists more information about a specific consumer's preference and consequently smaller variance.

a uniform prior distribution on ρ over a specified range. The parameter ρ must lie in the interval $[1/\lambda_{min}, 1/\lambda_{max}]$, where λ_{min} and λ_{max} denote the minimum and maximum eigenvalues of W for which the matrix $(I - \rho W)$ can be inverted (Sun, Tsutakawa, and Speckman 1999).

The Markov chain proceeds by generating draws from the set of conditional posterior distributions of the parameters. As mentioned previously, we augment the model parameters θ in Equation 7 that captured the dependent error structure through an autoregressive process. By conditioning on θ_i, the latent preference, z_i, arises from a standard binomial probit model with mean $x_i'\beta + \theta_i$ and independent errors. Furthermore, the conditional distributions of the model parameters, given θ, are of standard form. A detailed description of the full conditional distributions is provided in Appendix A. We note that generating draws from the full conditional distribution of θ is computationally demanding, and we adopted the method Smith and LeSage (2000) propose to generate draws iteratively for the elements of θ. Appendix B briefly outlines this method.

A SIMULATION STUDY

In this section, we demonstrate properties of the autoregressive choice model and investigate the relationship between sample size and accuracy of the Markov chain Monte Carlo estimator. We focus our analysis on the model without the latent mixing distribution described in Equations 11–13.

Data Simulation

We simulated three data sets. The first data set comprises 50 people, and the second data set comprises 500 people. We assume the people are connected circularly, with each person affected by his or her two closest neighbors. To illustrate, the autoregressive matrix W for five people is as follows:

$$W = \begin{bmatrix} 0 & .5 & 0 & 0 & .5 \\ .5 & 0 & .5 & 0 & 0 \\ 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & .5 & 0 & .5 \\ .5 & 0 & 0 & .5 & 0 \end{bmatrix}. \tag{18}$$

We assume that X comprises two covariates generated from a standard normal distribution, $\beta' = (1,1)'$, $\sigma^2 = 4$, and $\rho = .5$. We simulate binary choices (y_i) by generating draws from a multivariate normal distribution specified in Equation 10 and applying the censoring described in Equation 1. The covariance matrix of the latent preference distribution for the five people described by the autoregressive matrix in Equation 18 is

$$\Sigma = I + \sigma^2(I - \rho W)^{-1}(I - \rho W')^{-1} \tag{19}$$

$$= \begin{bmatrix} 7.3 & 3.2 & 1.6 & 1.6 & 3.2 \\ 3.2 & 7.3 & 3.2 & 1.6 & 1.6 \\ 1.6 & 3.2 & 7.3 & 3.2 & 1.6 \\ 1.6 & 1.6 & 3.2 & 7.3 & 3.2 \\ 3.2 & 1.6 & 1.6 & 3.2 & 7.3 \end{bmatrix}.$$

The covariance between the first and the third person is nonzero, despite that these two people are not neighbors. The first person is connected to the second person and the second to the third in the autoregressive matrix (Equation 18). The connections induce nonzero covariance between the first and the third person, reflecting the correlation of the circular network. The variances along the diagonal of Σ are equal, which reflects each person having exactly two neighbors. To examine the performance of the estimator for a heteroskedastic error matrix, we include a third case in which we formulated the matrix W using data from our empirical study (reported subsequently). This third simulation study contains 666 observations.

Accuracy of Parameter Estimates

Estimation results are presented in Table 1. We ran the Markov chain for 5000 iterations and deleted the first 1000 draws for "burn-in" of the chain. We used the last 4000 draws to calculate the posterior mean and standard deviation of the parameters. As Table 1 shows, the coefficient estimates are close to their true values, even in small samples. As we expected, the true parameters lie inside the 95% highest posterior density intervals of the posterior distributions, and the accuracy of the estimator improves as the sample size increases. We also ran the Markov chain for 10,000 iterations and found no appreciable difference in the estimates reported in Table 1. These results support our conclusion that the autoregressive choice model is statistically identified and that our estimation using data augmentation is valid.

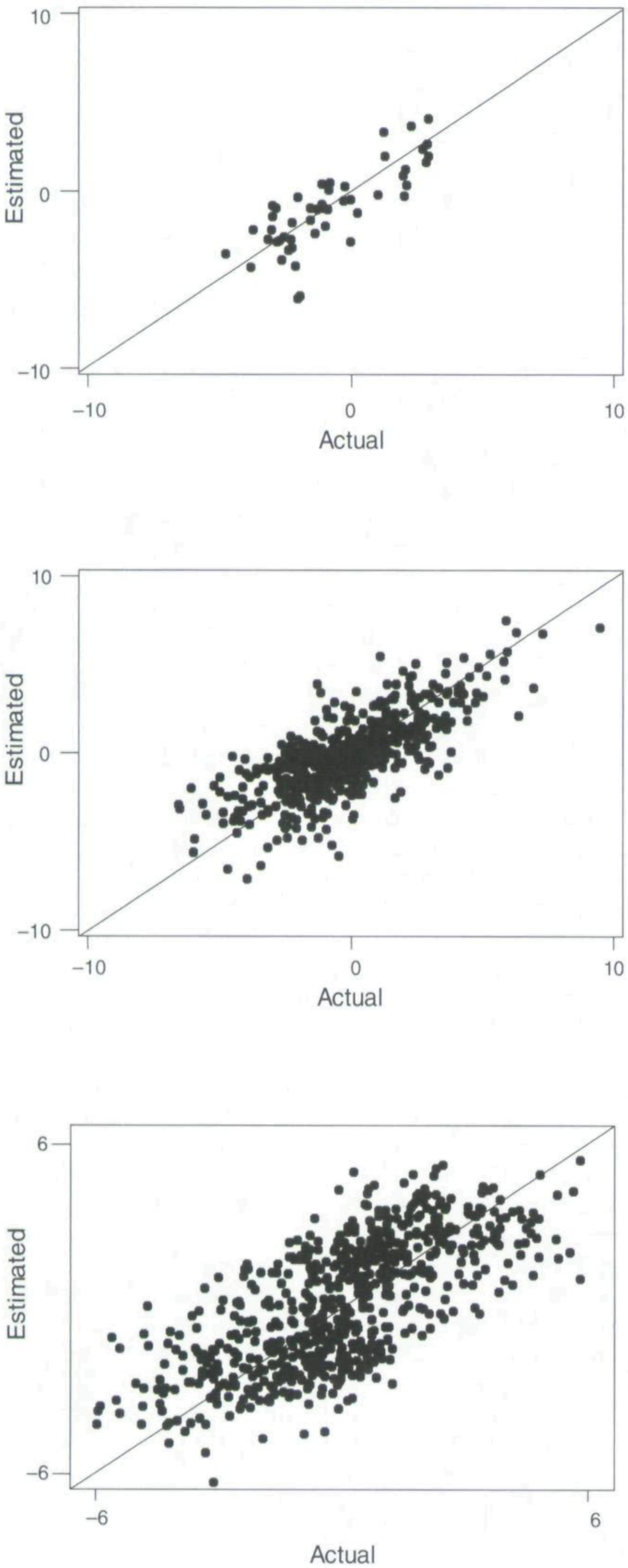
The purpose of the autoregressive specification is to understand dependence among consumers as reflected in the covariance matrix. It is therefore important to investigate the model's ability to recover the realizations of the choice model error responsible for inducing the covariances. Figure 1 provides a comparison between the estimated and the actual autoregressive effects (θ_i). The top panel of Figure 1

Table 1

ESTIMATES FROM THE NUMERICAL SIMULATION

			β_1	β_2	ρ	σ^2
N	*W*	*True Values*	*1*	*1*	*.5*	*4*
50	From simulation	Posterior mean	1.286	.884	.608	3.372
		Posterior standard deviation	(.622)	(.501)	(.089)	(1.859)
500	From simulation	Posterior mean	.951	.891	.510	4.010
		Posterior standard deviation	(.281)	(.290)	(.061)	(1.944)
666	From real data	Posterior mean	.921	.946	.474	3.250
		Posterior standard deviation	(.153)	(.163)	(.058)	(1.215)

Figure 1
ACTUAL VERSUS POSTERIOR ESTIMATES OF θ IN THE NUMERICAL SIMULATION



displays a plot of the estimated and actual autoregressive effects for the data set with 50 observations (people), the middle panel displays the plot for the data set with 500 observations, and the bottom panel displays the plot for the data set with 666 observations, with interaction generated to mimic our empirical analysis. Also included in the plots is a 45-degree line. The points in the graphs tend to fall evenly about the line, which indicates that the estimated autoregressive effects are not biased. Moreover, the variability of the points about the line does not depend appreciably on the sample size. An increase in the sample size results in improved estimates of the model parameters in Table 2 but not in terms of the augmented parameter θ_i, because the dimension of θ is the same as the number of observations in the analysis. Our analysis indicates that despite the sparseness present in a circularly connected population of 500 people or in our reported data, accurate estimates of model parameters, including the augmented parameters θ, are possible.

We use the autoregressive effects to assess the degree of dependence in the probit error structure. If there is little interdependence in consumer choice or if the data are not sufficiently informative about the presence of interdependencies, the realizations of θ_i will be near zero and the predicted choice probabilities, conditional on θ_i, will be similar to those obtained from an independent probit specification ($\Sigma = I$). If choice and preferences are interdependent, the realizations of θ_i will be large and improvements will exist in model fit by conditioning on the dependent information contained in θ_i.

Assessing Differences in Regression Coefficients

The presence of an autoregressive component in the error term changes the variance of the probit model from the identity matrix to $\Sigma = I + \sigma^2(I - \rho W)^{-1}(I - \rho W')^{-1}$. As illustrated in Equation 19, the autoregressive specification can lead to significant changes in the diagonal elements of the covariance matrix. Therefore, care must be exercised when interpreting the regression coefficients associated with the mean of the multivariate normal distribution. Although we have demonstrated that the autoregressive specification leads to an identified model, these coefficients must be interpreted relative to the scale of the error term. This is most easily accomplished by computing the expected change in the choice probability for a change in the independent variable.

Table 2 contains the regression coefficient estimates for the data set with 500 observations and regression estimates from a traditional binary probit model with $\Sigma = I$ (i.e., setting all θ_i to 0). The regression estimates for the autoregressive model are much larger than those obtained from the independent probit model; however, when these coefficients are converted to expected derivatives of the choice probability at $x_1 = 0$ and $x_2 = 0$, the estimates for the two models closely agree. The change in probability is .245 for the independent probit model (Equation 4) and .270 for the dependent probit model (Equation 10) when both x_1 and x_2 increase from .0 to 1. Either model is capable of capturing the average association between the covariates and choice; however, the autoregressive model is needed to understand the extent and nature of preference interdependencies given the average association.

Table 2
DIFFERENCES IN REGRESSION COEFFICIENTS

Model	True Values	β_1 1	β_2 1	ρ .5	σ^2 4
Independent	Posterior mean	.350	.317	0	0
(Equation 4)	Posterior standard deviation	(.058)	(.064)		
Dependent	Posterior mean	.951	.891	.510	4.01
(Equation 10)	Posterior standard deviation	(.281)	(.290)	(.061)	(1.644)

EMPIRICAL APPLICATION

A marketing research company collected data on purchases of midsize cars in the United States. We obscure the brand of the cars for the purpose of confidentiality. The cars are functionally substitutable, priced in a similar range, and distinguished primarily by their national origin: Japanese and non-Japanese. During the past 20 years, Japanese cars have acquired a reputation for reliability and quality, and we attempted to understand the extent to which preferences are interdependent among consumers.

We investigated two sources of dependence: geographic and demographic neighbors. Geographic neighbors are defined by physical proximity and are measured in terms of geographic distance among places of residence. Demographic neighbors are defined in terms of similar demographic variables. For example, young people are more likely to associate with other young people, obtain information from them, and want to conform to the beliefs of their reference group to gain group acceptance and social identity. We empirically tested these referencing structures and analyzed their importance in driving preferences.

We operationalized the different referencing schemes as follows: The data include information on the longitude and latitude information of each person's residence, and we can calculate the geographic distance between person i and person j as

$$d(i, j) = \sqrt{(d_i^1 - d_j^1)^2 + (d_i^2 - d_j^2)^2}, \tag{20}$$

where d_j^1 denotes the longitude and d_j^2 denotes the latitude of person j's home. We further assume that geographic influence is an inverse function of the geographic distance:

$$w_{i,j}^{g1} = \frac{1}{\exp[d(i, j)]}. \tag{21}$$

An alternative geographic specification of W that leads to a symmetric matrix is to identify neighbors by the Zip code of their home mailing address:

$$w_{i,j}^{g2} = 1 \text{ if person i and person j have the same Zip code; } 0 \text{ otherwise.} \tag{22}$$

We operationalized demographic neighbors in terms of people who share characteristics such as education, age, and income. We divided people in the data set into groups defined by age of the head of the household (three categories), annual household income (three categories), ethnic affiliation (two categories), and education (two categories), which leads to a maximum of 36 groups, 31 of which we present in our sample. The demographic specification of W becomes

$$w_{i,j}^{d} = 1 \text{ if person i and person j belong to the same demographic group; 0 otherwise.} \tag{23}$$

The data comprise 857 consumers who live in 122 different Zip codes. Table 3 provides sample statistics of the data. Approximately 85% of people in the sample purchased a Japanese car. On average, the price of a Japanese car is \$2,400 cheaper than a non-Japanese car, and there is little difference between the optional accessories purchased with the cars. However, we note that the sample standard deviations are nonzero, which indicates intragroup variation. The average age of the consumer is 49 years, and average annual household income is approximately \$67,000. Approximately 12% of the consumers are of Asian origin, and 35% have earned a college degree. We used the choices of 666 people from 100 Zip codes to calibrate the model, and 191 observations form a holdout sample. Figure 2 is a histogram of the number of Zip codes containing at least two consumers, which indicates that the sample of respondents is geographically dispersed.

We assessed in-sample and out-of-sample fit. When possible, we report fit statistics conditional on the augmented parameter θ_i. For the independent probit model (Equations 1–3) in which $\theta_i = 0$, the latent preferences are independent after we account for the influence of the covariates in mean of the latent distribution (Equation 4). Knowledge that a person actually purchased a Japanese car provides no help in predicting the preferences and choices of other people. When preferences are interdependent, information about others' choices is useful in predicting choices, and this information is provided through θ_i, as described in Equation 7.

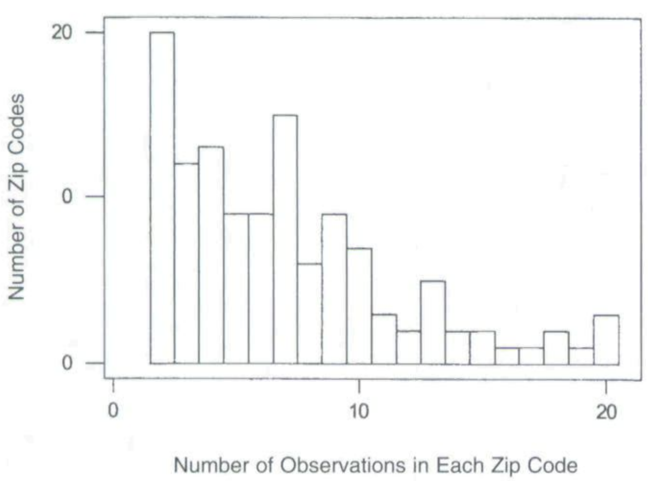
We assessed in-sample fit using Newton and Raftery's (1994, p. 21) importance sampling method that reweights the conditional likelihood of the data. Conditional on θ, this evaluation involves the product of independent probit probabilities and is easy to compute. Computation of the out-of-sample fit is more complicated. Our analysis proceeded as it would in developing a customer scoring model, by constructing the autocorrelation matrices W for the entire data set (857 observations) but by estimating the model parameters using the first 666 observations. We obtain the augmented parameters for the holdout sample, θ^p, by noting that

$$\begin{pmatrix} \theta \\ \theta^p \end{pmatrix} \approx N\left[0, \hat{\sigma}^2 (I_{857} - \hat{\rho}\hat{W})^{-1} (I_{857} - \hat{\rho}\hat{W}')^{-1}\right] = N\left(0, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right), \tag{24}$$

Table 3
SAMPLE STATISTICS

Variable	*Mean*	*Standard Deviation*
Car choice (1 = Japanese, 0 = non-Japanese)	.856	.351
Difference price (in $100)	–2.422	2.998
Difference in options (in $100)	.038	.342
Age of buyer (in number of years)	48.762	13.856
Annual income of buyer (in $1,000)	66.906	25.928
Ethnic origin (1 = Asian, 0 = non-Asian)	.117	.321
Education (1 = college, 0 = below college)	.349	.477
Latitude (relative to 30 = original –30)	3.968	.484
Longitude (relative to –110 = original 110)	–8.071	.503

Figure 2
HISTOGRAM OF SAMPLE SIZE FOR EACH ZIP CODE



where Σ_{12} is the covariance matrix between θ and θ^p, and we can obtain the conditional distribution of θ^p, given θ, using properties of the multivariate normal distribution:

$$(25) \qquad (\theta^p|\theta) \sim MN(\mu, \Omega),$$

where

$$\mu = \Sigma_{21}\Sigma_{11}^{-1}\theta, \text{ and}$$
$$\Omega = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}.$$

Table 4 reports the in-sample and out-of-sample fit statistics for six different models. The first model is an independent binary probit model (Equations 1–3) in which the probability of purchasing a Japanese car is associated with the feature differences among cars, individual demographic information, geographic information (longitude, latitude), and dummy variables for the demographic groups. The dummy variables are an attempt to capture high-order interactions of the covariates. We included dummy variables for only 19 of the 31 groups because the proportion of buyers of Japanese cars in the remaining groups is at or near 100%. Thus, the first model attempts to approximate the structure of heterogeneity using a flexible exogenous specification.

The second model is a random-effects model that assumes people who live in the same Zip code have identical price and option coefficients. We incorporated geographic and demographic variables into the model specification to adjust the model intercept and the mean of the random-effects distribution. The second model represents a standard approach to modeling preferences, incorporating observed and unobserved heterogeneity.

Models 3–6 specify four variations of interdependent models in which consumer preferences are endogenous, or interdependent. Models 3 and 4 are alternative specifications for geographic neighbors (Equations 21 and 22), whereas Model 5 specifies the autoregressive matrix in terms of the 31 demographic groups (Equation 23). Model 6 incorporates both geographic and demographic structures using the finite mixture model in Equations 11–13.

The model fit statistics (both in-sample and out-of-sample) indicate the following results. First, car choices are interdependent. Model 1 is the worst-fitting model, and all attempts to incorporate geographic and/or demographic interdependence in the model leads to improved in-sample and out-of-sample fit. A comparison of the fit statistics between Model 2 and Models 3–6 indicates that there is stronger evidence of interdependence in the autoregressive models than in a random-effects model based on Zip codes. Second, the addition of quadratic and cubic terms for longitude and latitude results in a slight improvement in in-sample fit (i.e., from –203.809 to –193.692) and out-of-sample fit (i.e., from .158 to .154). This supports the view that people have similar preferences not only because they share similar demographic characteristics that may point to similar patterns of resource allocation (an exogenous explanation) but also because there is an endogenous interdependence among people. Third, Models 3 and 4 produce similar fit statistics, which indicates that a geographic neighbor–based weighting matrix performs a good approximation to a geographic distance–based weighting matrix. The introduction of both geographic and demographic referencing schemes in Model 6 leads to an improvement in the fit and shows that both reference groups are important in influencing a person's preference.

Parameter estimates for the six models are reported in Table 5. We note that, in general, the coefficient estimates are largely consistent across the six models, which indicates that all the models are somewhat successful in reflecting the data structure. Because Model 6 yields the best in-sample and out-of-sample fit, we focus our discussion on its parameter estimates. The estimates indicate that price, age, income, ethnic origin, longitude, and latitude are significantly associated with car purchases; Asians, younger people, people

Table 4
MODEL FIT COMPARISON

Model	*Specification*[a]	*In-Sample Fit*[b]	*Out-of-Sample Fit*[c]
1 Standard probit model	$z_i = x_i^{1'}\beta^1 + x_i^{2'}\beta^2 + x_i^{3'}\beta^3 + x_i^{4'}\beta^4 + \varepsilon_i.$	–237.681	.177
2 Random coefficients model	$z_{ik} = x_i^{1'}\beta_k^1 + x_i^{2'}\beta^2 + x_i^{3'}\beta^3 + x_i^{4'}\beta^4 + \varepsilon_{ik},$ $\beta_k^1 = \Gamma(x_k^2, x_k^3) + \eta_k.$	–203.809	.158
3 Autoregressive model with geographic neighboring effect	$z_i = x_i^{1'}\beta^1 + x_i^{2'}\beta^2 + x_i^{3'}\beta^3 + x_i^{4'}\beta^4 + \theta_i + \varepsilon_i,$ $\theta_i = \rho\sum_{j=1}^{m} w_{ij}^{g1}\theta_j + u_i.$	–146.452	.136
4 Autoregressive model with geographic neighboring effect	$z_i = x_i^{1'}\beta^1 + x_i^{2'}\beta^2 + x_i^{3'}\beta^3 + x_i^{4'}\beta^4 + \theta_i + \varepsilon_i,$ $\theta_i = \rho\sum_{j=1}^{m} w_{ij}^{g2}\theta_j + u_i.$	–148.504	.135
5 Autoregressive model with demographic neighboring effect	$z_i = x_i^{1'}\beta^1 + x_i^{2'}\beta^2 + x_i^{3'}\beta^3 + x_i^{4'}\beta^4 + \theta_i + \varepsilon_i,$ $\theta_i = \rho\sum_{j=1}^{m} w_{ij}^{d}\theta_j + u_i.$	–151.237	.139
6 Autoregressive model with geographic and demographic neighboring effect	$z_i = x_i^{1'}\beta^1 + x_i^{2'}\beta^2 + x_i^{3'}\beta^3 + x_i^{4'}\beta^4 + \theta_i + \varepsilon_i,$ $\theta_i = \rho\sum_{j=1}^{m} w_{ij}\theta_j + u_i,$ $w_{ij} = \phi w_{ij}^{g2} + (1-\phi)w_{ij}^{d}.$	–133.836	.127

[a]$y_i = 1$ if a Japanese-made car is purchased (i indexes person and k indexes Zip code)
$x_i^1 = [1$, difference in price, difference in option$]$
$x_i^2 = [$age, income, ethnic group, education$]$
$x_i^3 = [$longitude, latitude$]$
$x_i^4 = [$group1, group2, ..., group19$]$
Note that for other subgroups, estimation of these group-level effects are not feasible because of the perfect or close-to-perfect correlation between the group dummy variable and y.
[b]We measured in-sample model fit by the log marginal density calculated using Newton and Raftery's (1994, p. 21) importance sampling method.
[c]We measured our sample fit by mean absolute deviation of estimated choice probability and actual choice.

with high incomes, and people who live relatively south and west prefer Japanese cars. The ethnic variable has a large coefficient, which indicates that Asian people have a significantly greater likelihood of choosing a Japanese car. Furthermore, ρ is significantly positive, which indicates a positive correlation among consumer preferences. In addition, α is significantly positive (ϕ greater than .5), which indicates that geographic reference groups are more important than demographic reference groups in determining individual preference.

Figure 3 displays estimates of the elements of the augmented parameter θ against the longitude and latitude of each observation in sample. Most of the estimates have posterior distributions away from zero, which provides evidence of interdependent choices. An analysis of the difference in covariates for the two groups ($\theta > 0$ and $\theta \leq 0$) does not reveal any statistically significant differences except for the longitude variable. That is, the augmented parameters that capture the endogenous nature of preferences are not simply associated with all the covariates in the analysis.

SUMMARY AND CONCLUSION

In this article, we introduce an autoregressive multivariate binomial probit model to study interdependent choices among consumers. We specify the model in a hierarchical Bayes framework, and we derive estimation algorithms using data augmentation to simplify the computations. We investigate the effects of two possible sources of interdependent influence: geographic neighbors and demographic neighbors. Geographic neighbors are people who reside in close proximity to one another, and demographic neighbors

Table 5
POSTERIOR MEAN OF COEFFICIENT ESTIMATES

Coefficient	*Model 1*	*Model 2*	*Model 3*	*Model 4*	*Model 5*	*Model 6*
Intercept	**–.799**	**1.015**	**1.135**	**1.489**	**1.116**	**1.274**
Price	**–.004**	**–.003**	**–.037**	**–.039**	**–.042**	**–.038**
Option	**.010**	**.005**	**.031**	**.034**	**–.028**	**–.015**
Age	**–.014**	**–.014**	**–.126**	**–.119**	**–.153**	**–.080**
Income	**.014**	**.001**	**.142**	**.146**	**.141**	**.167**
Ethnic	**2.271**	**2.042**	**7.513**	**7.679**	**7.382**	**9.670**
Education	**–.279**	**–.032**	**–3.662**	**–4.003**	**–3.914**	**–3.425**
Latitude	**–2.901**	**–5.778**	**–6.397**	**–6.965**	**–6.231**	**–6.841**
Longitude	**–1.739**	**–2.575**	**–3.216**	**–3.537**	**–3.330**	**–3.619**
Group1	–1.069	–1.214	–.897	–.765	–.682	–.552
Group2	–1.591	–1.123	–1.159	–1.264	–1.378	–1.432
Group3	–1.142	–1.256	–.569	–.512	–.165	–.273
Group4	–1.207	–.954	–3.159	–3.641	–2.331	–4.076
Group5	–.499	–.421	–2.315	–1.967	–1.578	2.523
Group6	–.703	–.946	–2.213	–2.786	–1.922	2.165
Group7	–.604	–.787	–1.887	–1.365	–2.134	2.632
Group8	–.238	–.113	1.459	1.781	.967	.957
Group9	–.445	–.335	2.646	2.893	1.384	2.484
Group10	–.597	–.667	.132	.197	.268	.169
Group11	**–1.207**	**–1.016**	**–3.376**	**–3.132**	**–2.711**	**–2.954**
Group12	**–1.634**	**–1.147**	**–4.105**	**–3.198**	**–3.452**	**–4.103**
Group13	–.333	–.178	2.139	2.164	2.551	2.872
Group14	**–1.669**	**–1.124**	2.778	2.115	2.160	2.949
Group15	–.630	–.421	1.986	1.468	1.235	1.836
Group16	**–.981**	**–.563**	**–2.071**	**–2.569**	**–2.135**	**–3.088**
Group17	–.577	–.289	1.468	1.923	1.497	2.313
Group18	–1.018	–.956	.912	.813	.569	.783
Group19	–.856	–.433	1.876	1.579	1.692	.964
σ^2	N.A.	N.A.	**16.784**	**16.319**	**15.691**	**18.235**
ρ	N.A.	N.A.	**.473**	**.462**	**.487**	**.512**
α	N.A.	N.A.	N.A.	N.A.	N.A.	**.598**

Notes: N.A. = not applicable. Bolded figures indicate that the 0 lies outside of the 95% highest posterior density interval of the estimate.

are people who share demographic variables that point to social networks.

We use alternative model specifications to investigate variation in preferences and find that a standard random-effect specification is inferior to an autoregressive specification. In the random-effect specification, we model variation in preferences across consumers as independent draws from a mixing distribution, whereas in an autoregressive specification the variation in preferences propagates through the networks. If person i is a neighbor of person j, and person j is a neighbor of person k, then person i can influence person k through person j. Such dependencies are not well reflected in i.i.d. draws from a mixing distribution.

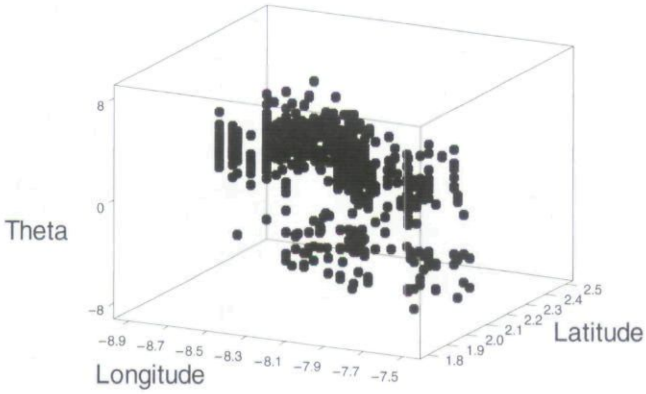
We apply the autoregressive model to a data set in which the dependent variable is whether an individual purchases a Japanese make of car. Our empirical application demonstrates that (1) there is a preference interdependence among individual consumers that reflects conformity ($\rho > 0$), (2) the preference interdependence is more likely to take an endogenous influence structure than a simple exogenous structure, and (3) the geographically defined network is more important than the demographic network in explaining individual consumer behavior. However, because our data are cross-sectional, we are unable to identify the true cause of the interdependence.

Choice models have been used extensively in the analysis of marketing data. In these applications, most of the analysis depends on the assumption that a consumer forms his or her own preferences and makes a choice decision irrespective of others' preferences. However, people live in a world in which they are interconnected, information is shared, recommendations are made, and social acceptance is important. Interdependence is therefore a more realistic assumption in models of preference heterogeneity.

Our model can be applied and extended in many ways. For example, opinion leaders exert a high degree of influence on others and could be identified with extreme realizations of the augmented parameter θ_i. Aspiration groups that affect others but are not themselves affected could be modeled with an autoregressive matrix W that is asymmetric. Temporal aspects of influence, including word of mouth and buzz (Rosen 2000), could be investigated, with longitudinal and cross-sectional data of the autoregressive matrix W defined on both dimensions. Such time-series data would help identify the source and nature of interdependence. Finally, the model can be extended to apply to multinomial response data to investigate the extent of interdependent preference in brand purchase behavior, or it can be extended to study the interdependence in β coefficients across consumers. These applications and extensions will contribute to the understanding of extended product offerings and the appropriateness of i.i.d. heterogeneity assumptions commonly made in models of consumer behavior.

Figure 3

ESTIMATED θ FOR EACH INDIVIDUAL (N = 666)



Profiling the Two Groups

Average	$\theta > 0$	$\theta \leq 0$	*p-Value*
Age	48.297	48.357	.962
Income	69.835	71.934	.418
Ethnic	.116	.094	.275
Education	.372	.379	.673
Latitude	4.152	4.158	.502
Longitude	–8.281	–8.186	.031

Notes: The origin for longitude and latitude is –110 and 30. All the longitude and latitude information presented here is related to this origin specification.

APPENDIX A: MARKOV CHAIN MONTE CARLO ESTIMATION

We carried out estimation by sequentially generating draws from the following distributions.

First, given the choice, a latent continuous variable z can be generated for the probit model. Generate $\{z_i, i = 1, \ldots, m\}$:

$$\text{(A1)} \qquad f(z_i|*) = \text{truncated normal}\left(x_{ij}{'}\beta + \theta_i, 1\right),$$

$$\text{if } y_i = 1, \text{ then } z_i \geq 0,$$

$$\text{if } y_i = 0, \text{ then } z_i < 0.$$

Second, generate β:

$$\text{(A2)} \qquad f(\beta|*) = MN(v, \Omega),$$

$$v = \Omega[X'(z - \theta) + D^{-1}\beta_0],$$

$$\Omega = (D^{-1} + X'X)^{-1},$$

$$\beta_0 = (0, 0, \ldots, 0)',$$

$$D = 400I_m.$$

Third, generate θ:

$$\text{(A3)} \qquad f(\theta|*) = MN(v, \Omega),$$

$$v = \Omega(z - X\beta),$$

$$\Omega = (I^{-1} + \sigma^{-2}B'B)^{-1},$$

$$B = I - \rho W.$$

Fourth, generate σ_γ^2:

$$\text{(A4)} \qquad f(\sigma_\gamma^2|*) \propto \text{inverted gamma}(a, b),$$

$$a = s_0 + m/2 \ (s_0 = 5),$$

$$b = \frac{2}{\theta'B'B\theta + 2/q_0} (q_0 = .1).$$

Fifth, generate ρ. We used the Metropolis-Hastings algorithm with a random walk chain to generate draws (see Chib and Greenberg 1995). Let $\rho^{(p)}$ denote the previous draw, and then the next draw $\rho^{(n)}$ is given by

$$\text{(A5)} \qquad \rho^{(n)} = \rho^{(p)} + \Delta,$$

with the accepting probability α given by

$$\min\left\{\frac{\left|B(\rho^{(n)})\right| \exp\left[-.5(1/\sigma^2)\theta'B(\rho^{(n)})'B(\rho^{(n)})\theta\right]}{\left|B(\rho^{(p)})\right| \exp\left[-.5(1/\sigma^2)\theta'B(\rho^{(p)})'B(\rho^{(p)})\theta\right]}, 1\right\}.$$

The Δ is a draw from the density normal(0, .005). The choice for parameters of this density ensures an acceptance rate of more than 50%. If ρ lies outside the range of $[1/\lambda_{min}, 1/\lambda_{max}]$, the likelihood is assumed to be zero, and we reject the candidate $\rho^{(n)}$.

Sixth, generate α:

$$\text{(A6)} \qquad W = \sum_{k=1}^{K} \phi_k W_k \text{ and } \phi_j = \frac{\exp(\alpha_j)}{\sum_{k=1}^{K} \exp(\alpha_k)}.$$

We elected to estimate α rather than ϕ directly. We used the Metropolis-Hastings algorithm with a random walk chain to generate draws (similar to generating ρ). Let $\alpha^{(p)}$ denote the previous draw, and then the next draw $\alpha^{(n)}$ is given by

$$\text{(A7)} \qquad \alpha^{(n)} = \alpha^{(p)} + \Delta,$$

with the accepting probability α given by

$$\min\Big(\Big\{\left|B(\alpha^{(n)})\right| \exp\left[-.5(1/\sigma^2)\theta'B(\alpha^{(n)})'B(\alpha^{(n)})\theta\right]$$

$$\exp\left[-.5(\alpha^{(n)} - \alpha_0)'T_o^{-1}(\alpha^{(n)} - \alpha_0)\right]\Big\}\Big/\Big\{\left|B(\alpha^{(p)})\right|$$

$$\exp\left[-.5(1/\sigma^2)\theta'B(\alpha^{(p)})'B(\alpha^{(p)})\theta\right]$$

$$\exp\left[-.5(\alpha^{(p)} - \alpha_0)'T_o^{-1}(\alpha^{(p)} - \alpha_0)\right]\Big\}, 1\Big).$$

The Δ is a draw from the density normal(0, .005I), α_0 is a vector of 0, and T_0 is a prior covariance matrix with diagonal elements equal to 100 and 0 for all off-diagonal elements. The choice for parameters of this density ensures an acceptance rate of more than 50%.

APPENDIX B: AN APPROACH TO EFFICIENTLY GENERATE θ

To generate θ, we needed to invert an mxm matrix $(I + \sigma^{-2}B'B)$. When m is large, it is computationally burdensome to make the inversion. We adopted an efficient method of generating θ from its posterior distribution while avoiding inverting a high-dimensional matrix that Smith and LeSage (2000) introduce. The essence is to generate univariate conditional posteriors of each component θ_i rather than a joint posterior distribution of the vector of θ. Next, we briefly describe the procedure:

$$\theta \propto \exp(-.5A),$$

where A

$$= \theta'(\sigma^{-2}B'B)\theta + (z - X\beta - \theta)'(z - X\beta - \theta),$$
$$= \sigma^{-2}\theta'(I - \rho W)'(I - \rho W)\theta + \theta'\theta - 2\phi'\theta,$$
$$= \sigma^{-2}(\theta'\theta - 2\rho\theta'W\theta + \rho^2\theta'W'W\theta) + \theta'\theta - 2\phi'\theta,$$

where

$$\phi = (\phi_i = z_i - x_i'\beta: i = 1, ..., m)'.$$

Further decompose the vector of θ as (θ_i, θ_{-i}), and define $w_{.i}$ to be the ith column of W and W_{-i} to be an mx (m – 1) matrix of all other columns of W. Then, we obtain the following expressions:

$$\theta'W\theta = \theta_i(\theta'w_{.i}) + \theta'W_{-i}\theta_{-i} = \theta_i \sum_{j \neq i} \theta_j(w_{ji} + w_{ij}) + C,$$

$$\theta'W'W\theta = \theta_i^2 w_{.i}'w_{.i} + 2\theta_i(w_{.i}'W_{-i}\theta_{-i}) + C,$$

$$\theta'\theta = \theta_i^2 + C,$$

$$-2\phi'\theta = -2\phi_i\theta_i + C,$$

where C denotes a constant that does not involve parameters of interest. Resubstituting the previous expressions enables us to write the conditional posterior distribution of θ_i as follows:

$$f(\theta_i|*) \propto \exp\left[-.5\left(a_i\theta_i^2 - 2b_i\theta_i\right)\right] = N\left(\frac{b_i}{a_i}, \frac{1}{a_i}\right),$$

where

$$a_i = \left(\frac{1}{\sigma^2}\right) + \left(\frac{\rho^2}{\sigma^2 w_{.i}'w_{.i}}\right) + 1,$$

$$b_i = \phi_i + \left[\frac{\rho}{\sigma^2 \sum_{j \neq i} \theta_j(w_{ji} + w_{ij})}\right] - \frac{\rho^2}{\left(\sigma^2 w_{.i}'W_{-i}\theta_{-i}\right)}.$$

We compared speed of this algorithm relative to generating θ by directly inverting the matrix $(I^{-1} + \sigma^{-2}B'B)$ and sampling from a multivariate distribution. The algorithm results in a fivefold decrease in the total time required for one iteration of the six-step algorithm for the second simulation study involving 500 observations.

REFERENCES

Alessie, Rob and Arie Kapteyn (1991), "Habit Formation, Interdependent Preferences, and Demographic Effects in the Almost Ideal Demand System," *The Economic Journal*, 101 (406), 404–419.

Allenby, Greg M. and Peter E. Rossi (1998), "Marketing Models of Marketing Models of Consumer Heterogeneity," *Journal of Econometrics*, 89 (1–2), 57–78.

Aronsson, Thomas, Sören Blomquist, and Hans Sacklén (1999), "Identifying Interdependent Behavior in an Empirical Model of Labor Supply," *Journal of Applied Econometrics*, 14 (6), 607–626.

Arora, Neeraj and Greg M. Allenby (1999), "Measuring the Influence of Individual Preference Structures in Group Decision Making," *Journal of Marketing Research*, 36 (November), 476–87.

Bearden, William O. and Michael J. Etzel (1982), "Reference Group Influence on Product and Brand Purchase Decisions," *Journal of Consumer Research*, 9 (3), 183–94.

Bernheim, Douglas B. (1994), "A Theory of Conformity," *Journal of Political Economy*, 10 (5), 841–77.

Bronnenberg, Bart J. and Vijay Mahajan (2001), "Unobserved Retailer Behavior I Multimarket Data: Joint Spatial Dependence in Marketing Shares and Promotion Variables," *Marketing Science*, 20 (3), 284–99.

——— and Catarina Sismeiro (2002), "Using Multimarket Data to Predict Brand Performance in Markets for Which No or Poor Data Exist," *Journal of Marketing Research*, 39 (February), 1–17.

Burnkrant, Robert E. and Alain Cousineau (1975), "Informational and Normative Social Influence in Buyer Behavior," *Journal of Consumer Research*, 2 (3), 206–215.

Case, Anne C. (1991), "Spatial Patterns in Household Demand," *Econometrica*, 59 (4), 953–65.

Chib, Siddhartha and Edward Greenberg (1995), "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, 49 (5), 327–35.

Childers, Terry L. and Akshay R. Rao (1992), "The Influence of Familial and Peer-Based Reference Groups on Consumer Decision," *Journal of Consumer Research*, 19 (2), 198–211.

Cowan, Robin, William Coman, and Peter Swann (1997), "A Model of Demand with Interactions Among Consumers," *International Journal of Industrial Organization*, 15 (6), 711–32.

Cressie, Noel A.C. (1991), *Statistics for Spatial Data*. New York: John Wiley & Sons.

Darrough, Masako, Robert A. Pollak, and Terence J. Wales (1983), "Dynamic and Stochastic Structure: An Analysis of Three Time Series of Household Budget Shares," *The Review of Economics and Statistics*, 66 (2), 274–81.

Duesenberry, James S. (1949), *Income, Saving, and the Theory of Consumer Behavior*. Cambridge, MA: Harvard University Press.

Duncan, O. Dudley, Archibald O. Haller, and Alejandro Portes (1968), "Peer Influences on Aspirations: Reinterpretation," *American Journal of Sociology*, 74 (2), 119–37.

Hayakawa, Hiroaki and Yiannis Venieris (1977), "Consumer Interdependence via Reference Groups," *Journal of Political Economy*, 85 (3), 599–615.

Hofstede, Frankel ter, Michel Wedel, and Jan-Benedict Steenkamp (2002), "Identifying Spatial Segments in International Markets," *Marketing Science*, 21 (2), 160–77.

Kamakura, Wagner A. and Gary J. Russell (1989), "A Probabilistic Choice Model for Marketing Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26 (November), 379–90.

Kapteyn, Arie (1977), "A Theory of Preference Formation," doctoral dissertation, Economics Department, Leyden University.

———, Sara van de Geer, Huib van de Stadt, and Tom Wansbeek (1997), "Interdependent Preferences: An Econometric Analysis," *Journal of Applied Econometrics*, 12 (5), 665–86.

Kelman, Herbert C. (1961), "Processes of Opinion Change," *Public Opinion Quarterly*, 25 (1), 57–78.

Leibenstein, Harvey (1950), "Bandwagon, Snob, and Veblen Effects in the Theory of Consumer's Demand," *Quarterly Journal of Economics*, 64 (2), 183–207.

LeSage, James P. (2000), "Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models," *Geographical Analysis*, 32 (1), 19–35.

Lessig, V. Parker and C. Whan Park (1978), "Promotional Perspective of Reference Group Influence: Advertising Implications," *Journal of Advertising*, 7 (2), 41–47.

——— and ——— (1982), "Motivational Reference Group Influence: Relationship to Product Complexity, Conspicuousness, and Brand Distinction," *European Research*, 10 (2), 91–101.

Manski, Charles F. (2000), "Economic Analysis of Social Interactions," *Journal of Economic Perspectives*, 14 (3), 115–36.

Newton, Michael A. and Adrian E. Raftery (1994), "Approximating Bayesian Inference with the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society (B)*, 56 (1), 3–48.

Pollak, Robert A. (1976), "Interdependent Preference," *American Economic Review*, 66 (3), 309–320.

Rosen, Emanuel (2000), *The Anatomy of Buzz*. New York: Doubleday.

Smith, Tony E. and James P. LeSage (2000), "A Bayesian Probit Model with Spatial Dependencies," working paper, Department of Systems Engineering, University of Pennsylvania.

Sun, Dongchu, Robert K. Tsutakawa, and Paul L. Speckman (1999), "Posterior Distribution of Hierarchical Models Using car(1) Distribution," *Biometrika*, 86(2), 341–50.

Tanner, Martin A. and Wing Hung Wong (1987), "The Calculation of Posterior Distributions by Data Augmentation," *Journal of the American Statistical Association*, 82 (398), 528–40.

van de Stadt, Huib, Arie Kapteyn, and Sara van de Geer (1985), "The Relativity of Utility: Evidence from Panel Data," *The Review of Economics and Statistics*, 67 (2), 179–87.