The "Shopping Basket": A Model for Multicategory Purchase Incidence Decisions

Puneet Manchanda • Asim Ansari • Sunil Gupta

Graduate School of Business, University of Chicago, Chicago, Illinois 60637, puneet.manchanda@gsb.uchicago.edu
Graduate School of Business, Columbia University, New York, New York 10027, maa48@columbia.edu
Graduate School of Business, Columbia University, New York, New York 10027, sg37@columbia.edu

Abstract

Consumers make multicategory decisions in a variety of contexts such as choice of multiple categories during a shopping trip or mail-order purchasing. The choice of one category may affect the selection of another category due to the complementary nature (e.g., cake mix and cake frosting) of the two categories. Alternatively, two categories may co-occur in a shopping basket not because they are complementary but because of similar purchase cycles (e.g., beer and diapers) or because of a host of other unobserved factors. While complementarity gives managers some control over consumers' buying behavior (e.g., a change in the price of cake mix could change the purchase probability of cake frosting), co-occurrence or co-incidence is less controllable. Other factors that may affect multi-category choice may be (unobserved) household preferences or (observed) household demographics. We also argue that not accounting for these three factors simultaneously could lead to erroneous inferences. We then develop a conceptual framework that incorporates complementarity, co-incidence and heterogeneity (both observed and unobserved) as the factors that could lead to multicategory choice.

We then translate this framework into a model of multicategory choice. Our model is based on random utility theory and allows for simultaneous, interdependent choice of many items. This model, the multi*variate* probit model, is implemented in a Hierarchical Bayes framework. The hierarchy consists of three levels. The first level captures the choice of items for the shopping basket during a shopping trip. The second level captures differences across households and the third level specifies the priors for the unknown parameters. We generalize some recent advances in Markov chain Monte Carlo methods in order to estimate the model. Specifically, we use a substitution sampler which incorporates techniques such as the Metropolis Hit-and-Run algorithm and the Gibbs Sampler.

The model is estimated on four categories (cake mix, cake frosting, fabric detergent and fabric softener) using multicategory panel data. The results disentangle the complementarity and co-incidence effects. The complementarity results show that pricing and promotional changes in one category affect purchase incidence in related product categories. In general, the cross-price and cross-promotion effects are smaller than the own-price and own-promotions effects. The cross-effects are also asymmetric across pairs of categories, i.e., related category pairs may be characterized as having a "primary" and a "secondary" category. Thus these results provide a more complete description of the effects of promotional changes by examining them both within and across categories. The co-incidence results show the extent of the relationship between categories that arises from uncontrollable and unobserved factors. These results are useful since they provide insights into a general structure of dependence relationships across categories. The heterogeneity results show that observed demographic factors such as family size influence the intrinsic category preference of households. Larger family sizes also tend to make households more price sensitive for both the primary and secondary categories. We find that price sensitivities across categories are not highly correlated at the household level. We also find some evidence that intrinsic preferences for cake mix and cake frosting are more closely related than preferences for fabric detergent and fabric softener.

We compare our model with a series of null models using both estimation and holdout samples. We show that both complementarity and co-incidence play a significant role in predicting multicategory choice. We also show how many single-category models used in conjunction may not be good predictors of joint choice.

Our results are likely to be of interest to retailers and manufacturers trying to optimize pricing and promotion strategies across many categories as well as in designing micromarketing strategies. We illustrate some of these benefits by carrying out an analysis which shows that the "true" impact of complementarity and co-incidence on profitability is significant in a retail setting. Our model can also be applied to other domains. The combination of item interdependence and individual household level estimates may be of particular interest to database marketers in building customized "cross-selling" strategies in the direct mail and financial service industries.
(*Multicategory Models; Shopping Baskets; Retailing; Micromarketing; Multivariate Probit Model; Hierarchical Bayes Models*)

1. Introduction

Consumers make multicategory decisions in a variety of contexts such as grocery shopping trips, mail-order purchasing, and financial portfolio choice. A common thread that underlies these multicategory choice scenarios is that the chosen items may be related to each other in some manner e.g., a pair of shoes may be ordered to go along with a particular dress. In the retailing context, these multicategory decisions result in the formation of consumers' "shopping baskets" which comprise the collection of categories that consumers purchase on a specific shopping trip.[1] Retailers are interested in understanding the composition of the shopping basket in terms of the cross-category dependencies among the purchased categories. Industry experts have suggested that the identification of "substitute" and "complementary" categories is critical to developing an understanding of how these shopping baskets are put together (Terbreek 1993). Besides the cross-category focus, retail managers are also interested in implementing micromarketing programs through the use of individual household level data (Rossi et al. 1996; *Food Marketing Institute* 1994; Mogelonsky 1994). A combination of cross-category insights and individual household-level preferences could potentially lead to retailers' partitioning the overall portfolio of retail categories into smaller independent sub-portfolios, e.g., dairy products and cleaning products may be managed as independent sub-portfolios. This would then allow retailers to make pricing and promotion decisions to optimize profits across these sub-portfolios for each individual household.

The insights developed from cross-category analysis coupled with micromarketing are also of interest to manufacturers and database marketers. Manufacturers, such as P & G, who market brands in related categories, could utilize these insights to rationalize marketing expenditures across two or more categories, e.g., the toothpaste and toothbrush categories. Database marketers are interested in discovering and converting cross-category dependencies to targeted cross-selling programs (Berry 1994). These programs

[1]We use the accepted notion of "category" as defined by managers and data providers such as IRI and Nielsen.

typically use knowledge about a specific consumer's intrinsic preferences or purchasing patterns in one category to induce purchase in the other, e.g., American Express cards may send a customer a coupon for golf clothing after observing a purchase of golf clubs.

In spite of significant industry interest in understanding cross-category purchase behavior of consumers, most research in marketing has been on single category choice decisions (see Meyer and Kahn 1991). This research has focused on independent analyses of consumer purchasing activities within categories without explicitly capturing the interdependence across categories. It is only recently that researchers have started examining purchases across multiple categories (Russell et al. 1997).

A common industry practice to study cross-category purchasing behavior is to cross-tabulate joint purchases across multiple categories. However, this approach cannot distinguish among the possible mechanisms leading to joint purchasing activity. The choice of one category may affect the selection of another category due to the complementary nature (e.g., cake mix and cake frosting) of the two categories. Alternatively, two categories may co-occur in a shopping basket not because they are complementary but because of similar purchase cycles (e.g., beer and diapers) or because of a host of other unobserved factors (e.g., consumer movement through a store). While complementarity allows managers some control over consumers' buying behavior (e.g., a change in the price of cake mix could change the purchase probability of cake frosting), co-occurrence or co-incidence is less controllable.

In this research, we develop a framework for investigating multicategory purchase decisions in the grocery shopping context. This framework incorporates three crucial elements that could influence these choices—complementarity, co-incidence and heterogeneity. In more specific terms, this paper has three objectives. The first objective is to develop a general model of multicategory choice that explicitly allows for dependence across the chosen categories, separates the effects of controllable drivers of multicategory choice from the uncontrollable drivers, and accounts for household specific preferences for each category. The second objective is to test the model in the context of consumers' shopping baskets. The final objective is to

provide retailers and manufacturers insights into the nature of the relationship between categories.

We translate our framework into a Hierarchical Bayes model of multicategory purchases. Specifically, we use a multivariate probit model and estimate it using Markov chain Monte Carlo techniques such as the Metropolis-Hastings Algorithm and the Gibbs Sampler. The paper, therefore, makes both substantive and methodological contributions.

The structure of the paper is as follows. We review the previous work in this area in §2. Section 3 deals with the conceptual background which motivates our modeling approach. Section 4 describes the model structure. In §5 we use our approach to develop the model of the shopping basket. Specifically, we discuss the data, model specification, variable operationalization, estimation procedure, the results and the comparison with null models in this section. Section 6 discusses the managerial implications of our findings. We conclude the paper with directions for future research in §7.

2. Previous Research

Previous modeling research in the multicategory domain can be classified into three streams.[2] The first stream of research examines a phenomenon of interest (say, promotion elasticity) across many categories and "meta-analyzes" the results to obtain empirical generalizations. One of the earliest papers in this stream is by Fader and Lodish (1990) who showed that certain consumer characteristics such as household penetration and frequency of purchase tend to explain the pricing and promotional environment in retail stores at the category level. An extension of this research by Narasimhan et al. (1996) showed that promotional elasticity of a category depends on category structure and consumer characteristics. An earlier paper by Raju (1992) examined the variability in category sales and

[2]Mention must also be made of other related work in an experimental setting that examines behavior across multiple categories. Menon and Kahn (1995) examine variety seeking behavior and show that changes in the choice context (a different category) can affect behavior in the target category. Ratneshwar et al. (1996) show that, given goal directed behavior, consumers form consideration sets across product categories.

linked it to both category characteristics (e.g., bulkiness) and marketing variables (e.g., promotion). Finally, Hoch et al. (1995) relate the store-level price elasticities of various categories to the demographics of consumers in the trading areas of the stores. Although this stream of research generalizes across many categories, it ignores interdependence in consumers' purchases across multiple categories—which is the focus of our research.

The second stream of research examines how a dependent variable of interest (e.g., store choice) is affected by multicategory variables. For example, Bell and Lattin (1998) examine how store choice is affected by the pricing strategy of a store across many categories. Though they do not specifically model dependence across categories, they allow for a multicategory flavor through the use of a composite variable (total amount spent on a shopping trip). Similarly, Bodapati (1996) looks at the impact of feature advertising across many categories on consumers' store choice decision. Again, this stream of research does not explicitly capture interdependence in consumer purchase behavior across categories.

The third stream of research *explicitly* allows for dependency across multicategory items. Walters (1991) and Mulhern and Leone (1991) model store sales using regression methods and show that sales of brands in one category are affected by the pricing and promotion of brands in a predefined "complementary" category. Chintagunta and Haldar (1998) examine purchase timing across predefined complement and substitute category pairs and show that the nature of the relationship between category pairs influences interpurchase times. Manchanda and Gupta (1997) examine purchase incidence and brand choice for pairs of categories by creating composite basket alternatives. A major limitation of these approaches is that, while they account for complementarity, they ignore either co-incidence or heterogeneity or both. As we will highlight in the next section, this omission may lead to erroneous inferences.

The focus of our paper is to build a general framework that incorporates aspects of complementarity, co-incidence and heterogeneity in modeling multicategory consumer purchase decisions.

3. Conceptual Background

A pair of categories can be bought together on a shopping trip by a variety of reasons:

1. Purchase Complementarity or Cross-Effects: Two categories may be complementary. In such a situation, we can expect the marketing activity (price or promotion) in one category to influence consumers' purchase of the other category. For example, a price discount on beer may cause the consumer to buy beer as well as potato chips. We define two categories as purchase complements if marketing actions (price and promotion) in one category influence the purchase decision in the other category.[3] We do this to separate the effects of controllable and uncontrollable factors. Price and promotion are within managers' control while most of the other factors are either not controllable (e.g., consumer heterogeneity) or not observed by the researcher.

2. Co-incidence: We define co-incidence as the set of all reasons except purchase complementarity and consumer heterogeneity that could induce joint purchase of items across categories. For example, consumers may shop for many items on the same shopping trip because of economic reasons, (e.g., consumers can spread the cost of a trip over many items), or owing to habit (Kahn and Schmittlein 1992). Similarly, other unobserved factors such as consumer knowledge of a particular supermarket, time pressure on the consumer during the trip (Park et al. 1989) or consumer mood (Donovan et al. 1994) may influence joint purchasing. The physical environment of the store may also induce joint purchases (Swinyard 1993). Co-incidence refers to all the "residual" reasons that could cause joint incidence after accounting for the impact of marketing variables and heterogeneity.

3. Heterogeneity: Finally, consumer heterogeneity may also result in aggregate data that shows high joint purchasing between two categories. Accounting for heterogeneity may be crucial in understanding the true nature of the association across categories. We postulate that households differ in their intrinsic utilities and

price responses for purchase incidence in each category and that these utilities and responses may be related across categories. We also allow for the fact that household demographics may explain some part of the relationship among the intrinsic household utilities and price responses.

It is crucial to account for the impact of both marketing variables (i.e., cross-effect or complementarity) and the unobserved factors (co-incidence) on cross-category purchasing activity. Ignoring either of these sets of factors could lead to erroneous inferences. To illustrate this, we examine the relationship between a pair of categories on both dimensions. In Table 1, on the complementarity dimension, purchases between a pair of categories may be negatively related (e.g., two categories may be substitutes or "negative complements"), independent, or positively related (e.g., price promotion in one category may increase the probability of purchase incidence in the other category). Similarly, a pair of categories may be related positively on the co-incidence dimension, i.e., they may be bought together due to similar purchase cycles, they may be independent, or they may be negatively related. The observed joint occurrences of purchases in a pair of categories is the net effect of these two factors. For example, when both co-incidence and complementarity are positive, the observed data will show very strong evidence of joint purchases (cell 9). However, when two categories have positive complementarity but negative co-incidence (cell 3) or vice versa (cell 7), it is not clear whether the net effect will be positive, zero, or negative.

In general, two types of misleading inferences may arise from ignoring either co-incidence or complementarity. In the first type, we might infer complementarity or co-incidence when it was not actually present.

[3]Note that this definition applies to two categories being purchased together as distinct from two categories being consumed together. We do not model consumption complementarity as we do not observe it.

Table 1 Co-incidence and Complementarity Scenarios

Co-incidence	Complementarity		
	−	0	+
−	− −[1]	−[2]	+/−[3]
0	−[4]	0[5]	+[6]
+	+/−[7]	+[8]	+ +[9]

This could happen if cell 8 represents the "true" model but we erroneously infer cell 6 to be the represented behavior. In the second case, cell 3 may represent the true model, but if the " + " complementarity and " − " co-incidence effects cancel out, we may erroneously conclude that cell 5 is the correct specification. In other words, we may infer complementarity when it is not there (first case) or we may infer no complementarity when it truly exists (second case).

4. Modeling Approach

We build our model in three stages. In the first stage we model complementarity (or cross-effects) and co-incidence for a single household. In the second stage, we account for heterogeneity across households by specifying a population distribution over the household specific parameters. In the third stage, we specify priors over the parameters of the population distribution.

4.1. Stage 1

We observe each household, h, making purchase incidence decisions across a set of J categories on a store trip t. These decisions can be represented by a vector $\mathbf{y}_{ht} = \{y_{h1t}, y_{h2t}, \ldots, y_{hJt}\}$, of binary dependent variables. In keeping with a random utility formulation, we model this observed behavior in terms of latent utilities for the categories. Specifically, the underlying utilities for the J categories can be written as

$$u_{h1t} = \beta_{h10} + \beta_{h11} \text{ Own effects}$$
$$+ \beta_{h12} \text{ Cross effects} + \epsilon_{h1t}$$
$$u_{h2t} = \beta_{h20} + \beta_{h21} \text{ Own effects}$$
$$+ \beta_{h22} \text{ Cross effects} + \epsilon_{h2t}$$
$$\vdots$$
$$u_{hJt} = \beta_{hJ0} + \beta_{hJ1} \text{ Own effects}$$
$$+ \beta_{hJ2} \text{ Cross effects} + \epsilon_{hJt}.$$

The link between the observed behavior and the latent utility for any category j can be represented as follows:

$$y_{hjt} = \begin{cases} 1, & \text{if } u_{hjt} > 0, \\ 0, & \text{if } u_{hjt} \leq 0. \end{cases}$$

In this utility specification, the cross-effects capture the impact of purchase complementarity. The coefficients for the cross-effects contained in β_{hj2} represent the change in purchase utility of category j due to the marketing actions of other categories. The direct impact of marketing actions of the same category is captured by the coefficients of the own effects β_{hj1}.

The utility equations for household h on store trip t for the J categories can be compactly represented as:

$$u_{ht} = X_{ht}\beta_h + \epsilon_{ht}, \tag{1}$$

where the jth row of the matrix X_{ht} contains the causal variables influencing the utility for the jth category and $\beta_h = \{\beta_{h0}, \beta_{h1}, \beta_{h2}\}$ contains the household specific coefficients.

The unobserved influences operating at the shopping trip are represented by $\epsilon_{ht} = \{\epsilon_{h1t}, \epsilon_{h2t}, \ldots, \epsilon_{hJt}\}$. As these unobserved factors may be common across categories, we assume that

$$\epsilon_{ht} \sim \text{MVN } [0, \Sigma], \tag{2}$$

where Σ is a $J \times J$ covariance matrix. This correlated error structure of the purchase utilities captures co-incidence. Thus if cov $(\epsilon_{hit}, \epsilon_{hjt}) > 0$, then an increase in the purchase utility of category i will lead to an increase in the purchase utility of category j. In other words, the error correlations capture the linkages between the uncontrollable factors that drive joint purchases.[4]

The above formulation results in a multi*variate* probit model (Greene 1997, Chib and Greenberg 1998). The multivariate probit model is particularly suited for our investigation as it allows for more than one category to be purchased simultaneously. This is an extremely important feature since it allows us to model both the size of the basket (how many items) and the composition of the basket (which items). It is important to note that the multivariate probit model is distinct from the multi*nomial* probit model (McCulloch and Rossi 1994) which only allows the choice of one alternative from a set of mutually exclusive alternatives.

---

[4]Note that the co-incidence term arises from a combination of these residual reasons as well as biases due to model misspecification (such as omitted variables or not accounting for consumer consideration sets). This must be kept in mind whenever an attempt is made to draw inferences on the basis of the error correlation terms. We thank the Area Editor for bringing this to our attention.

The binary nature of the observed dependent variables necessitates the introduction of appropriate scaling constants to ensure identifiability of model parameters. In the multivariate probit model, the components of the observed incidence profile, y_{ht}, remain unaltered if each of the corresponding latent utilities in u_{ht} is multiplied by a (different) positive constant. We can therefore fix the scale of the utilities by dividing each utility by its corresponding standard deviation, thereby yielding an identified set of parameters. The identification restrictions transform Σ to a correlation matrix. The signs of the correlations in Σ provide insights about the qualitative nature of co-incidence. The magnitudes provide a measure of the strength of the impact of unobserved factors in inducing joint purchasing activity. In our subsequent discussion we will retain the same set of symbols, $(\beta, \Sigma \text{ and } u)$, to denote identified parameters and utilities.[5]

4.2. Stage 2
In the second stage we model the observed and unobserved sources of heterogeneity across households by specifying a population (mixing) distribution over the household specific coefficients. The household level parameters can therefore be formulated as a linear function of demographic variables, i.e.,

$$\beta_h = D_h \mu + \lambda_h, \quad h = 1 \text{ to } H, \qquad (3)$$

where $\beta_h = \{\beta_{h0}, \beta_{h1}, \beta_{h2}\}$, $\lambda_h \sim MVN[0, \Lambda]$, and D_h is a matrix containing the demographic and other household specific variables.[6] The vector, μ, captures the impact of demographic variables and thus provides insights about the observed sources of heterogeneity across households. The J dimensional vector, $\lambda_h = \{\lambda_{h1}, \lambda_{h2}, \ldots, \lambda_{hJ}\}$, represents the unobserved sources of heterogeneity across households. The variance-covariance matrix, Λ, represents the residual covariation in the household specific parameters after having adjusted for the variation due to observed sources of heterogeneity.

The multivariate probit model as specified above is

---

[5]Further details on the identification for each set of parameters are available from the authors.

[6]This is the most general specification of heterogeneity. In our application, we impose a specific structure on the Λ matrix in order to reduce model complexity. This is discussed later.

empirically intractable. Greene (1997, pp. 911–912) points out that models with more than three dimensions are extremely difficult to estimate as it is hard (computationally) to evaluate the high order multivariate normal integrals that are required in specifying the likelihood. Another source of difficulty arises in ensuring a positive definite correlation matrix for the errors. We circumvent the need for high dimensional integration by using numerical Bayesian approaches that build upon recent advances in Markov chain Monte Carlo (MCMC) methods (Chib and Greenberg 1998, Dey and Chen 1996). We use a hierarchical Bayesian approach, which is necessary for modeling observed and unobserved sources of consumer heterogeneity. Some examples of hierarchical models in the marketing literature include Allenby and Ginter (1995), Rossi et al. (1996) and Ansari et al. (1996). Since the Bayesian approach requires priors over unknown quantities in the model, the third stage of our model specifies these priors.

4.3. Stage 3
We provide a brief description of the priors in this section (details are given in appendix A1). We specify independent and diffuse (but proper) priors over the population level parameters, Σ, μ, and Λ^{-1}. Specifically, the prior for μ is multivariate normal, $MVN[0, \Theta]$, and the prior for the precision matrix, Λ^{-1}, is Wishart, $W[\rho, (\rho R)^{-1}]$. Finally, we assume that the nonredundant correlation parameters, $\text{vec}^*(\Sigma) = (\sigma_{12}, \sigma_{13}, \ldots \sigma_{(J-1)J})'$, of the Σ matrix come from a truncated multivariate normal population with a fixed mean and covariance matrix (Chib and Greenberg 1998).

4.4. Inference Procedure
The probability of observing an incidence profile, $y_{ht} = \{y_{h1t}, \ldots, y_{hJt}\}$, on a single observation is given by

$$\Pr(Y_{ht} = y_{ht} \mid \beta_h, \Sigma) = \int_{s1} \cdots \int_{sJ}$$

$$\frac{1}{\sqrt{(2\pi) \det(\Sigma)^{1/2}}} \exp\left(-\frac{1}{2} \epsilon'_{ht} \Sigma^{-1} \epsilon_{ht}\right) d\epsilon_{ht}, \qquad (4)$$

where $\epsilon_{ht} = u_{ht} - X_{ht}\beta_h$, and

$$S_j = \begin{cases} (-\infty, 0), & \text{if } y_{hjt} = 0, \\ (0, \infty), & \text{if } y_{hjt} = 1. \end{cases} \qquad (5)$$

The likelihood for the household h is given by

$$\prod_{t=1}^{n_h} \Pr(y_{ht} \mid \beta_h, \Sigma), \qquad (6)$$

and the unconditional likelihood for an arbitrary household is obtained as

$$\int \int \cdots \int \prod_{t=1}^{n_h} \Pr(y_{ht} \mid \beta_h, \Sigma) \, f(\beta_h) d\beta_h. \qquad (7)$$

This likelihood is clearly very complicated as it involves computation of high order multidimensional integrals making classical inference based on maximum likelihood methods difficult. In the Bayesian framework, inference about the unknown parameters is based on their joint posterior distribution. We use simulation based methods to make many random draws from this posterior distribution. Inference is then based on the empirical distribution of this sample of draws. MCMC methods are used to simulate the draws. Specifically, we use substitution sampling (a combination of the Gibbs sampler and the Metropolis Hastings Hit-and-Run Algorithm) in tandem with data augmentation (Albert and Chib 1993) to obtain a sample of parameter draws from the joint posterior distribution.

Substitution sampling replaces one complicated draw from the joint posterior with a sequence of relatively simple draws from easy to sample full conditional distributions of parameter blocks (Appendix A2 details the full conditional distributions for our model). Since the full conditional distribution for some parameter blocks is not known in our application, we cannot use a Gibbs sampling step for those parameters. We therefore replace it by a Metropolis-Hastings step (Hastings 1970). We generalize methods reported in Chib and Greenberg (1998) and Dey and Chen (1996) to provide an MCMC algorithm for a hierarchical Bayes multivariate probit specification. In the context of the model developed in the previous section, the $(m + 1)$th step of the substitution sampling algorithm involves generating the following draws:

a. Generate β_h draws from $p(\beta_h^{(m+1)} \mid \{\{u_{ht}^{(m)}\}\}, \Sigma^{(m)}, \mu^{(m)}, \Lambda^{(m)})$ for $h = 1$ to H.

b. Generate a μ draw from $p(\mu^{(m+1)} \mid \{\beta_h^{(m+1)}\}, \Lambda^{(m)}, 0, \Theta)$.

c. Generate a Λ^{-1} draw from $p(\Lambda^{-1(m+1)} \mid \{\beta_h^{(m+1)}\}, \{\mu^{(m+1)}\}, \rho, R)$.

d. Generate u_{ht} draws from $p(u_{hjt} \mid \beta_h^{(m+1)} \Sigma^{(m)}, y_{ht})$ for $h = 1$ to H, $t = 1$ to n_h.

e. Generate a Σ matrix from $p(\Sigma^{(m+1)} \mid \{\{u_{ht}^{(m+1)}\}\}, \{\beta_h^{(m+1)}\}, \Sigma_0, G_0)$ using a Metropolis-Hastings Hit-and-Run step.

This sequence of draws generates a Markov chain whose stationary distribution is the joint posterior density of all unknowns. The initial output from the chain reflects a transient or "burn-in" period in which the chain has not converged to the equilibrium distribution and is therefore discarded. A sample of draws obtained after convergence is used to make posterior inferences about model parameters and other quantities of interest.

5. The Shopping Basket

In this section we describe the modeling of the shopping basket in the context of the proposed approach. We describe the data, the model specification and variable operationalization, null models, and the estimation procedure. We then discuss the results of the proposed model and compare these with the results from the null models. We conclude the section with details on model prediction.

5.1. Data

The data were drawn from a multicategory dataset made available by A. C. Nielsen. The data span a period of 120 weeks from January 1993 to March 1995 and are from a large metropolitan market area in the western United States.

Four grocery categories were used for this application—Laundry Detergent (Det), Fabric Softener (Soft), Cake Mix (Cake) and Cake Frosting (Frost). The categories were chosen to illustrate some of the features of our proposed model and to check the face validity of some of our results. For example, we expect that detergent-softener and cake mix-frosting pairs to be (purchase) complements while the remaining pairs are likely to be (purchase) independent.

Households that had a minimum of two purchases in each of the four categories during the 120 week period were selected. This resulted in a sample of 205 households. We then randomly split this sample into two parts. The first part consisted of 155 households

(estimation sample) and the second part consisted of 50 households (prediction sample).

The total number of purchase occasions for households in the estimation sample was 17,389. Out of these at least one category was purchased on 3,414 occasions and no purchase was made on the remaining 13,975 occasions. The number of pairwise purchases (for all six category pairs) across the 3,414 occasions is detailed in Table 2. An examination of this table shows high dependence between detergent and softener and between cake mix and frosting and little dependence across the other pairs. The purpose of our model is not only to validate these obvious pairings but to also assess the relative effects of complementarity (cross-effects), co-incidence, and heterogeneity.

5.2. Model Specification and Variable Definition

The deterministic component of the utility for category j for household h at time t can be written as

$$u_{hjt} = \text{Constant}_{hjt} + \text{Own Effects}_{hjt}$$
$$+ \text{Cross Effects}_{hjt}. \qquad (8)$$

Direct Impact or Own Effects: These effects for category j are specified as[7]

$$\text{Own Effects}_{hjt} = \beta_{h1j} * \text{Price}_{hjt} + \gamma_{1j} * \text{Promo}_{hjt}. \qquad (9)$$

The Price and Promotion variables were constructed as follows:

[7]There is a tradition of using household inventory as an explanatory variable to predict purchase incidence in the marketing literature (e.g., Bucklin and Gupta 1992). However, the use of this variable necessitates some strong assumptions such as the use of a linear, constant rate of consumption and raises concerns about endogeneity. We therefore do not include the inventory variable in the model specification reported here though we estimated models with and without inventory variables.

Price: We used a "category price" variable (in $/Oz.) for each household on each shopping trip. This was done by constructing the category price for each household as a weighted average price of brands where the weights were the share of each brand bought by each household.[8]

Promotion: A brand was considered to be on promotion if it was either featured or displayed. The "category promotion" variable was constructed in a similar manner to that of the category price variable. This resulted in a variable that took a value between 0 and 1. Some descriptive statistics for the four categories are given in Tables 3a–3c.

Complementarity or Cross Effects: We model the cross-effects of all other categories on category j using a similar set of variables.

$$\text{Cross Effects}_{hjkt} = \sum_k \beta_{h2jk} * \text{Price}_{hkt}$$
$$+ \sum_k \gamma_{1jk} * \text{Promo}_{hkt}, \qquad (10)$$

where $k = 1, \ldots, J$ and $k \neq j$.

Heterogeneity: The full model has 36 variables. A completely unstructured representation of heterogeneity over the coefficients for these variables would necessitate estimating 666 unique parameters of the 36×36 covariance matrix, Λ. In order to reduce model complexity, we use a structured representation of heterogeneity. Specifically, we assume that the promotion

[8]Since our model is not based on an axiomatic framework, it does not prescribe a "correct" category price operationalization. In other words, a weighted average household level price is just one of many possible alternatives. We chose this operationalization since, (a) it allows for household level differences by constructing a price based on a household's past purchases and (b) it has been used in past research (e.g., Krishnamurthi and Raj 1988). We also estimated models with different operationalizations, e.g., price paid was used as category price when a brand was purchased and a weighted average otherwise. We found no directional differences among the results and the estimated price elasticities were marginally different in magnitude. Our modeling approach is in contrast to some other model formulations where a specific error structure results in a "category" level variable e.g., the Category Value term in a nested logit setting when the errors are distributed i.i.d Gumbel. In our future research we plan to investigate model structures that do not allow ad hoc definitions of the price variable while allowing for flexible error structures.

Table 2 Frequencies of Joint (Pairwise) Incidence

	Det	Soft	Cake	Frost
Det	1201	275	40	24
Soft		593	27	11
Cake			415	610
Frost				219

variables have coefficients that are invariant across households.[9]

We capture observed sources of heterogeneity by allowing household-specific base utilities, β_{h0}, the own price coefficients, β_{h1}, and the cross price coefficients, β_{h2}, to be a function of demographic variables. We use two demographic variables, Family Size (FS) and Total Trips (TT). We hypothesize that larger households (with more children) may consume cake mix and cake frosting more often and may consume larger quantities of laundry products—this would be reflected in more frequent purchases of these categories. The mean family size in the data was 3.2 and the range was from 1 to 8. Similarly, there is some evidence that households who shop more often are different in their response characteristics than those that shop less often (Ainslie and Rossi 1998).[10] We aggregated the total number of shopping trips (across a large number of categories) for each household to construct this variable. The average number of trips for these households in our data was 139 with a range from 26 to 317. Therefore, the household specific parameters can be written as follows:

$$\beta_{h0j} = \mu_{01j} + \mu_{02j} * FS_h + \mu_{03j} * TT_h + \lambda_{h0j}, \quad (11)$$

$$\beta_{h1j} = \mu_{11j} + \mu_{12j} * FS_h + \mu_{13j} * TT_h + \lambda_{h1j}, \quad (12)$$

$$\beta_{h2j} = \mu_{21j} + \mu_{22j} * FS_h + \mu_{23j} * TT_h + \lambda_{h2j}, \quad (13)$$

where λ_{h0j}, λ_{h1j} and λ_{h2j} are the error terms that represent the residual unobserved heterogeneity after accounting for observed heterogeneity.

The above specification of the utility equations for the J categories can be summarized as:

$$u_{ht} = \beta_{h0} + X_{h1t}\beta_{h1} + X_{h2t}\beta_{h2} + Z_{ht}\gamma + \epsilon_{ht}, \quad (14)$$

$$\beta_{h0} = D_h\mu_0 + \lambda_{h0}, \quad (15)$$

$$\beta_{h1} = D_h\mu_1 + \lambda_{h1}, \quad (16)$$

[9]Our finding across many model specifications was that own and cross promotions effects were less important than the own and cross price effects. Hence, in the interest of model parsimony, we chose to keep these coefficients, γ, invariant across households. The prior for γ is multivariate normal, MVN [η, C]. γ draws are included as a step in the substitution sampler described earlier. More details are given in Appendices 1 and 2.

[10]We thank one of the reviewers for bringing this to our attention.

Table 3a — Summary Statistics

Category	Mean # of Purchases	# of Brands	Mean Price ($/Oz.)	Mean Promotion
Det	10	30	0.064	0.05
Soft	6	7	0.054	0.06
Cake	8	4	0.063	0.19
Frost	6	4	0.103	0.10

Table 3b — Price Correlation Across Categories

	Det	Soft	Cake	Frost
Det	1.00	0.20	0.03	0.09
Soft		1.00	0.09	0.11
Cake			1.00	0.15
Frost				1.00

Table 3c — Promotion Correlation Across Categories

	Det	Soft	Cake	Frost
Det	1.00	0.04	0.01	0.03
Soft		1.00	0.04	0.05
Cake			1.00	0.22
Frost				1.00

$$\beta_{h2} = D_h\mu_2 + \lambda_{h2}, \quad (17)$$

where $\epsilon_{ht} \sim N(0, \Sigma)$, $\lambda_{h0} \sim N(0, \Lambda_0)$, $\lambda_{h1} \sim N(0, \Lambda_1)$, and $\lambda_{h2} \sim N(0, \Lambda_2)$. The matrix X_{h1t} is a $J \times J$ matrix containing the own price variables for the J categories, the matrix X_{h2t} is a $J \times J$ matrix containing the cross price variables, and Z_{ht} contains the own and cross promotion variables.

5.3. Null Models

We compare our model with three null models (note that all null models incorporate heterogeneity):

Null Model 1 (Independent Probits): The simplest null model is one that accounts for neither co-incidence nor complementarity. This implies that the deterministic part of the model contains only own effects. This model also assumes that the errors are drawn from a multivariate normal distribution with $\mathbf{0}$ mean and an identity variance-covariance matrix. This is identical to estimating a series of J independent hierarchical probit models, one for each category.

Null Model 2 (Independent Probits with Cross-Effects): In the next model, we add complementarity to null model 1 but still do not account for co-incidence. This is achieved by adding the cross price and promotion terms to the deterministic part of the utility.

Null Model 3 (Multivariate Probit with Co-Incidence): We then add co-incidence to null model 1 but do not add cross effects. This is achieved by allowing the variance-covariance matrix of the error terms to be an unrestricted correlation matrix.

This selection of null models would help in understanding the effects of complementarity and co-incidence alone as well as together.

5.4. Estimation

The models were estimated using C programs developed by the authors. As explained in § 4, repeated draws were made from the series of full conditionals to arrive at the joint posterior density of the unknown quantities using the MCMC sampling scheme. The substitution sampler was run for 50,000 iterations for the proposed model and null models. Convergence was ensured by monitoring the time-series of the draws from the full conditional distributions (cf. Allenby and Ginter 1995). As indicated earlier, initial iterations reflect a "burn-in" period where the chain may not have converged. We chose a burn-in length of 45,000 iterations. Therefore, we retained only the last 5,000 draws from the posterior distributions for inference purposes. Since sequential draws from the joint posterior may be highly correlated, we resort to "thinning the chain" (Geyer 1992, Raftery and Lewis 1995) whereby we retain every fifth draw from the undiscarded part of the chain. Therefore, 1,000 draws from the posterior distribution of each parameter were retained and used to make inferences.

5.5. Results

5.5.1. Proposed Model. Results from the proposed model are given in Tables 4a–4f.[11] As is usual

[11] As mentioned earlier, we estimated models with and without own and cross inventory variables. The inclusion or exclusion of these variables did not affect the value of the other parameters in the model. For models that included the inventory variables, we found results that were consistent with our expectation. Own inventory

for Bayesian inference, we summarize the posterior distribution of the parameters by reporting the posterior mean and 95% probability interval.[12]

Own Effects: The own price effects (Table 4a) are in the expected direction. The household specific price coefficients are negative and significant for all four categories, i.e., a reduction in category price increases the probability of incidence for that particular category.[13] There is a significant, positive effect of promotions for laundry detergent and fabric softener but not for cake mix and cake frosting (Table 4b). This suggests that perhaps cake mix and frosting purchases are more planned than the purchases of detergent and softener even though the promotion frequency for these categories is higher than that of detergent and softener.

Cross-Effects: We first estimated a model where the cross effects of all three other categories were included for a particular category. However, as expected, we found that the only significant set of cross-effects were for detergent-softener and cake mix-cake frosting pairs. No cross-effects were found across any other pair of categories. In the interest of reduced model complexity and faster convergence we set the nonsignificant cross-effects to zero and reestimated the model.[14] The results from this model are reported in Tables 4a–4b.

Four interesting results emerge from these tables. First, we find that all four cross price effects, i.e., softener on detergent, detergent on softener, cake frosting on cake mix and cake mix on cake frosting, are negative and significant. In other words, a decrease in cake frosting price increases the probability of buying cake

effects were negative and significant for all four categories. The cross inventory coefficients were positive and significant for the detergent-softener and cake mix-frosting pairs. This indicates that consumption complementarity may also play a role in inducing joint purchases.

[12] In our subsequent discussion we refer to a variable as "not significant" if 0 is included in the probability interval of its coefficient.

[13] We report the mean (across all households) and the range for the household specific price parameters (both own and cross) in Table 4a. More details on the β_h parameters are provided when we discuss heterogeneity.

[14] There was no significant difference between the estimates from the full and the reduced model.

Table 4a   Direct Impact and Complementarity (Price) Parameter Estimates

Price of	Incidence Category			
	Det	Soft	Cake	Frost
Det	-25.75^a $(-30.98, -10.30)^b$	-7.20^a $(-14.90, -0.21)^b$		
Soft	-3.28^a $(-8.08, -0.36)^b$	-10.81^a $(-20.42, -6.82)^b$		
Cake			-17.16^a $(-30.96, -10.30)^b$	-13.42^a $(-29.59, -0.88)^b$
Frost			-8.10^a $(-18.32, -1.86)^b$	-11.70^a $(-20.89, -5.75)^b$

^{a}Mean across all households.

^{b}Range across all households.

mix. The implications of this finding are significant for retailers and managers who manage brands across categories. Retailers now have an additional marketing variable that they can use to influence category incidence and basket composition while managers can optimize promotion budgets across brands in different categories. Second, the coefficients of the cross-price effect are smaller than the coefficients of the own-price effect. Third, we see a strong asymmetry in cross-price effects, e.g., the effect of a change in detergent price on softener purchase is much larger than the other way around. Since the comparison of the "raw" price coefficients provides a rough estimate of the cross-price effects, we defer a detailed discussion of this asymmetry to a later section when we discuss own and cross-price elasticities. Fourth, although we expected the sign of the cross-promotion coefficients to be positive among complements, we find no significant effects except for the effect of detergent promotion on softener purchase. Once again, it shows the asymmetry between the effects of promotion across the two categories.

Co-incidence: The co-incidence pattern as described by the correlation matrix of errors (Table 4c) shows some interesting results. As expected, the correlation coefficients between cake mix and cake frosting and detergent and softener are high (0.92 and 0.46 respectively), when compared to the correlations between

other pairs. The difference between the magnitudes of the two correlations mentioned above could be because cake mix and cake frosting are purchased together more often in our data set than detergent and softener. This might be due to the fact that cake mix and cake frosting have higher consumption complementarity than detergent and softener. Also, as mentioned previously, the nonsignificant effects of promotions for cake mix and cake frosting suggest that perhaps the purchase of these two categories is planned leading to high co-incidence. Of the other four category pairs, two of the correlation coefficients (softener-cake mix and softener-frosting) are not significantly different from 0, and the remaining two are about 0.05 reflecting a very low level of co-incidence.

Category Specific Intercepts (β_{h0j}): We first discuss the mean estimates of the category specific intercepts (Table 4d). We find that larger families tend to increase the mean utility of purchase for the cake mix and frosting categories. This may result from the fact that the primary consumers of prepared cakes in a household are children. The total number of trips made by each household do not seem to have a significant effect on the mean purchase utility for any category.

The covariance matrix, Λ_0, of the residual mean category specific utilities provides insights into the nature of household heterogeneity across these four categories. A transformation of the covariance matrix, Λ_0,

into a correlation matrix yields additional insights. The mean utilities for detergent and softener have a correlation coefficient of 0.24 while those of cake mix and cake frosting have a correlation coefficient of 0.58. This suggests that, after accounting for family size and total trips, there exist some more household specific unobserved variables that link the utilities of these two pairs of categories at the *individual household* level. One possible hypothesis is that these correlations reflect the extent of consumption complementarity amongst the two pairs of categories. The fact that all the other four correlations are near 0 provides some additional support for this hypothesis. Further, the correlation between cake mix and frosting is larger than the correlation between detergent and softener, once again suggesting stronger complementarity between cake mix and frosting than between detergent and softener.

We now examine the effect of demographics on the own and cross price coefficients (Tables 4e–4f). We find evidence that price sensitivity is affected by both household demographics and household shopping behavior. For own price, the coefficients for family size and total trips are negative and significant across all four categories. This implies that large families are more price sensitive and that families that make more frequent shopping trips are also more price sensitive. Similarly, larger families tend to be more price sensitive as far as cross price sensitivity is concerned, though total trips do not seem to affect this. These findings are consistent with those reported in Ainslie and Rossi (1998).

In summary, we find that both observed and unobserved sources of heterogeneity (such as the demographic and shopping variables used here) play a part in explaining household response. Family size seems to be a better predictor of price response than the total number of trips (10 out of 12 coefficients are affected by family size while 4 out 12 are affected by total number of trips).

Overall, to put these results in the context of Table 1, we find that the detergent and softener and cake and frosting pairs occupy cell 9 (positive complementarity and positive co-incidence) while all the other pairs seem to be from cell 5 (independent in both complementarity and co-incidence).

5.5.2. Elasticity Analysis. There is no analytical

Table 4b **Direct Impact and Complementarity (Promotion) Parameter Estimates and 95% Probability Intervals[1]**

Promotion of	Incidence Category			
	Det	Soft	Cake	Frost
Det	**0.75** (0.53, 0.98)	**0.18** (0.02, 0.35)		
Soft	−0.02 (−0.20, 0.15)	**0.35** (0.06, 0.61)		
Cake			−0.01 (−0.13, 0.11)	−0.05 (−0.15, 0.05)
Frost			−0.07 (−0.26, 0.08)	−0.14 (−0.33, 0.10)

[1]Significant estimates are denoted in boldface type.

Table 4c **Co-incidence Results (Correlation Matrix Σ) Parameter Estimates and 95% Probability Intervals**

	Det	Soft	Cake	Frost
Det	1.00	**0.46** (0.44, 0.47)	**0.05** (0.01, 0.08)	**0.04** (0.01, 0.07)
Soft		1.00	0.03 (−0.01, 0.07)	0.00 (−0.02, 0.02)
Cake			1.00	**0.92** (0.91, 0.93)
Frost				1.00

expression for determining the incidence elasticities for the multivariate probit model. We therefore use simulation to make 1,000 draws from a multivariate normal distribution using the estimated mean β's and the X's (e.g., price) and the mean Σ correlation matrix. The resulting utility measures are used to estimate the base incidence probabilities. We then make a 10% change in the specific independent variable and recalculate the incidence probabilities in a similar manner.[15] We calculate (purchase) incidence elasticities in the usual manner. The estimated elasticities are reported in Table 5. The own-price elasticities are different

[15]We also estimated the elasticities for a 5% and 15% change in the independent variable. There was no appreciable difference in the estimated elasticities in all three cases.

Table 4d   Heterogeneity Results (β_{h0j})

Category Intercepts

$$\beta_{h0j} = \mu_{01j} + \mu_{02j} * FS_h + \mu_{03j} * TT_h + \lambda_{h0j}$$

Cat	μ_{01}	μ_{02}	μ_{03}
Det	**0.69**	1.06	−0.04
	(0.18, 1.18)	(−0.51, 2.44)	(−0.06, 0.02)
Soft	**−1.06**	1.24	−0.00
	(−1.44, −0.57)	(−0.28, 2.76)	(−0.01, 0.02)
Cake	**−0.32**	**4.40**	−0.05
	(−0.65, −0.02)	(3.75, 5.05)	(−0.12, 0.02)
Frost	**−0.94**	**3.77**	0.01
	(−1.30, −0.58)	(2.66, 4.59)	(−0.02, 0.03)

Table 4e   Heterogeneity Results (β_{h1j})

Own Price Coefficients

$$\beta_{h1j} = \mu_{11j} + \mu_{12j} * FS_h + \mu_{13j} * TT_h + \lambda_{h1j}$$

Cat	μ_{11}	μ_{12}	μ_{13}
Det	**−16.88**	**−1.60**	**−0.03**
	(−19.67, −10.92)	(−2.97, −1.09)	(−0.05, −0.01)
Soft	**−4.62**	**−0.34**	**−0.04**
	(−5.81, −3.13)	(−0.58, −0.05)	(−0.06, −0.02)
Cake	**−9.29**	**−2.33**	**−0.01**
	(−10.92, −7.12)	(−2.61, −2.05)	(−0.03, −0.01)
Frost	**−2.53**	**−1.62**	**−0.03**
	(−6.32, −0.32)	(−2.11, −0.96)	(−0.05, −0.02)

Table 4f   Heterogeneity Results (β_{h2j})

Cross Price Coefficients

$$\beta_{h2j} = \mu_{21j} + \mu_{22j} * FS_h + \mu_{23j} * TT_h + \lambda_{h2j}$$

Cat	μ_{21}	μ_{22}	μ_{23}
Det/Soft[1]	**−0.61**	**−0.87**	−0.05
	(−1.33, −0.08)	(−3.81, −0.13)	(−0.23, 0.10)
Soft/Det	**−6.04**	**−0.44**	0.01
	(−11.55, −1.15)	(−0.98, −0.06)	(−0.08, 0.10)
Cake/Frost	**−1.81**	**−2.04**	−0.04
	(−3.22, −0.49)	(−3.34, −0.61)	(−0.36, 0.29)
Frost/Cake	**−4.69**	**−2.59**	0.14
	(−8.25, −0.88)	(−3.52, −0.57)	(−0.16, 0.41)

[1] An entry of Cat A/Cat B represents the effect on Cat A purchase due to a change in price of Cat B.

Table 5   Mean Price Elasticities

| Price of | Probability of Incidence in Category | | | |
	Det	Soft	Cake	Frost
Det	−0.40	−0.12		
	(0.14)[a]	(0.05)		
Soft	−0.06	−0.70		
	(0.02)	(0.30)		
Cake			−0.17	−0.15
			(0.04)	(0.06)
Frost			−0.11	−0.21
			(0.04)	(0.06)

[a]Numbers in parentheses represent posterior standard deviations.

across categories and range from −0.17 to −0.70. Within the two related pairs, detergent and cake mix have lower own price elasticities than softener and frosting, respectively.[16]

As expected, the magnitude of each cross-price elasticity is much lower than the magnitude of the associated own-price elasticity. The cross-price elasticities show an interesting asymmetric pattern. Price changes of detergent have a larger effect on softener purchase (cross price elasticity of −0.12) than the other way around (cross price elasticity of −0.06). Cake mix and cake frosting also show a similar pattern where price changes of cake mix have a slightly large effect on cake frosting purchase (cross price elasticity of −0.15) than vice versa (cross price elasticity of −0.11).

These results suggest that in each pair of categories, there may be one "primary" category that drives purchase incidence for the pair. These primary categories (detergent and cake mix) have lower own-price elasticities, but changes in their prices have a stronger effect on the purchase incidence of the associated "secondary" category. In other words, detergent and cake mix prices are stronger drivers of softener and cake frosting purchases than vice versa. These results seem to be consistent with our a priori knowledge.

The promotion elasticities are much smaller in magnitude than the price elasticities. The own-promotion

[16]Note that when we say one elasticity is lower than another, we are referring to the absolute magnitudes of the estimates.

elasticity for detergent is 0.017 while the own-promotion elasticity for softener is 0.005. The cross-promotion elasticity of detergent promotion on softener incidence is 0.003. Again, we find that the pattern of elasticities indicates detergent to be the primary category for this pair.

A major advantage of estimating household specific parameters is that it enables us to examine response across categories at the household level. One possible question could be whether each household is price sensitive to a similar extent across all categories. This question can be partly answered by converting the error covariance matrices of the own and cross price coefficients, Λ_1 and Λ_2, to correlation matrices, and examining the pairwise correlations. The error correlations for own price are: Det-Soft (0.18), Det-Cake (-0.09), Det-Frost (0.08), Soft-Cake (0.09), Soft-Frost (-0.07), Cake-Frost (0.25). The error correlations for cross price are: Det-Soft (0.11), Det-Cake (-0.04), Det-Frost (0.06), Soft-Cake (0.03), Soft-Frost (-0.06), Cake-Frost (0.19).

There are two interesting findings here. First, it seems that households do not have price sensitivities that are similar across categories (the own and cross price parameters show four positive correlations and two negative correlations). Second, the magnitude of the correlations is not very high (with higher correlations for the detergent-softener and cake mix-frosting pairs and lower correlations for the other four pairs). This suggests that price sensitivities may depend on a combination of category or category groups and household characteristics, e.g., a household may have similar price sensitivity for the cleaning products, which may be distinct from price sensitivity for baking products. Earlier studies have shown differing results. In the context of brand choice across unrelated categories, Kim and Srinivasan (1995) show that households do not seem to have similar price sensitivities. In a similar context, however, Ainslie and Rossi (1998) find evidence that household price sensitivities are somewhat related. We speculate that this mixed bag of results may arise from differences in the behavior being modeled (purchase incidence, conditional brand choice) or the specific method that is used. Further research is needed before any definitive conclusions can be drawn.

Another advantage of obtaining household level parameters is that these parameters may be translated into own-price and cross-price elasticities for all the 155 households in the sample. These elasticities could prove very useful in applications such as database marketing.

5.6. Comparison with Null Models

As mentioned earlier, we estimated three null models (Table 6). The proposed model is compared with these null models on price elasticities (Table 7), co-incidence parameters and prediction using both the estimation sample (Table 8) and a holdout sample (Table 9).

5.6.1. Price Elasticity and Co-incidence Comparison.

Price Elasticities: We compare the complementarity results across models by focusing on the own and cross-price elasticities. Since simpler model specifications (which omit complementarity or co-incidence) are likely to overstate the effect of price, we expect price elasticity to be the highest for null model 1 and the lowest for null model 4. Similarly, we expect that the cross-price elasticities would be higher for null model 2 than for the proposed model. Our hypothesis is that the change in elasticities is likely to be higher for the cross price elasticities due to the direct trade-off between complementarity and co-incidence.

The comparison of the elasticities is detailed in Table 7. For the own price elasticities, we observe a drop (ranging from 10% to 24%) in the magnitudes between null model 1 and the proposed model. There appears to be no difference between the estimates from null model 1, 2, and 3. Consistent with our hypothesis, the cross price elasticities show a larger drop in magnitude

Table 6   Proposed and Null Models

Model Accounts For	Null Model 1 (Independent Probits)	Null Model 2 (Independent Probits)	Null Model 3 (Multivariate Probit)	Proposed Model (Multivariate Probit)
Complementarity	No	Yes	No	Yes
Co-incidence	No	No	Yes	Yes
Heterogeneity	Yes	Yes	Yes	Yes

Table 7   Price Elasticity Comparison Across Models

Price Elasticity	Null Model 1[b]	Null Model 2[a]	Null Model 3[a]	Proposed Model
Own				
Det	−0.44	−0.43	−0.43	−0.40
	(0.18)[e]	(0.17)	(0.17)	(0.14)
Soft	−0.78	−0.76	−0.75	−0.70
	(0.34)	(0.34)	(0.31)	(0.30)
Cake	−0.21	−0.21	−0.20	−0.17
	(0.07)	(0.05)	(0.06)	(0.04)
Frost	−0.24	−0.23	−0.23	−0.21
	(0.07)	(0.07)	(0.07)	(0.06)
Cross[a]				
Det/Soft		−0.08		−0.06
		(0.03)		(0.02)
Soft/Det		−0.16		−0.12
		(0.07)		(0.05)
Cake/Frost		−0.21		−0.11
		(0.08)		(0.04)
Frost/Cake		−0.24		−0.15
		(0.11)		(0.06)

[a]An entry of Cat A/Cat B represents the effect on Cat A purchase due to a change in price of Cat B.

[b]Independent probit models.

[c]Independent probit models with complementarity.

[d]Multivariate probit model with co-incidence.

[e]Numbers in parentheses represent posterior standard deviations.

Table 8   Predicted Baskets on Estimation Sample

Basket Number	Basket Description[a]	Frequency	Null Model 1[b]	Null Model 2[c]	Null Model 3[d]	Proposed Model
1	N	13975	13387	13589	14239	13936
2	D	1201	1383	1319	1221	1250
3	S	593	724	697	612	570
4	C	[415]	[851]	[643]	[417]	[399]
5	F	219	596	519	210	207
6	DS	274	104	185	239	261
7	DC	40	93	53	37	38
8	DF	24	66	59	20	28
9	SC	20	62	55	29	35
10	SF	8	37	55	7	6
11	CF	[502]	[61]	[155]	[268]	[535]
12	DSC	7	9	15	10	13
13	DSF	3	5	14	3	4
14	DCF	60	7	6	44	64
15	SCF	30	5	24	23	29
16	DSCF	18	1	2	10	15
"Hit Rate"			0.85	0.90	0.96	0.99

[a]N implies that no category is purchased. D, S, C, F imply that there was a purchase in the detergent, softener, cake mix or frosting categories respectively.

[b]Independent probit models.

[c]Independent probit models with complementarity.

[d]Multivariate probit model with co-incidence.

(from 33% to 75%) between null model 2 and the proposed model. Interestingly, the pair with the higher co-incidence (cake mix-frosting) shows a bigger effect in the drop in cross price elasticity magnitude.

Co-incidence: We expect the error to decrease as the models increase in sophistication and explanatory power. In other words, we expect that the effect of co-incidence is reduced as we move from null model 3 to the proposed model.[17] For both the correlations that are of interest, i.e., detergent and softener and cake mix and frosting, we see some evidence of decline. The correlation for detergent and softener drops from 0.51 to 0.46, while for cake mix and frosting it declines from 0.95 to 0.91. These results are consistent with our expectations.

In conclusion, on the substantive front, the proposed

[17]Note that co-incidence is not modeled in null models 1 and 2.

model provides better diagnostics than the null models. We find that the addition of controllable (price and promotion) factors can reduce the extent of co-incidence for a category pair (e.g., detergent and softener).[18]

5.6.2. Model Comparison. We now compare models by focusing on the predictive ability of the models on both estimation sample of 155 households as well as a holdout sample of 50 households. We generate the probability of incidence for each category by using a frequency simulator based on 1,000 draws from the estimated model. We use these category incidence probabilities to generate the proportions for each of the 16 baskets, e.g., detergent incidence only

[18]In other analyses not reported in this paper, we found some evidence that accounting for uncontrollable factors such as unobserved heterogeneity also tends to reduce the size of the error correlation.

Table 9 Predicted Baskets on Holdout Sample

Basket Number	Basket Description[a]	Frequency	Null Model 1[b]	Null Model 2[c]	Null Model 3[d]	Proposed Model
1	N	3811	3707	3665	3686	3771
2	D	248	208	245	288	169
3	S	139	169	154	106	184
4	C	94	166	141	70	55
5	F	44	168	138	110	105
6	DS	37	12	55	53	47
7	DC	3	12	19	7	4
8	DF	0	12	18	9	7
9	SC	2	8	7	2	5
10	SF	0	8	7	3	4
11	CF	92	8	18	127	107
12	DSC	2	1	4	1	2
13	DSF	0	1	4	1	2
14	DCF	4	1	2	12	9
15	SCF	4	0	4	4	8
16	DSCF	1	0	0	2	3
"Hit Rate"			0.88	0.90	0.92	0.93

[a]Independent probit models.

[b]Independent probit models with complementarity.

[c]Multivariate probit model with co-incidence.

corresponds to basket 2. We then sum up the proportions across all the observations and translate the proportion into the number of baskets, e.g., an aggregate proportion of 0.86 translates to 3854 baskets (0.88 × 4481 where 4481 is the total number of baskets). The predicted baskets from each model are detailed in Table 8 (estimation sample) and Table 9 (holdout sample). We then define a simple "Hit Rate" metric to compare these aggregate level predictions across models as follows:

$$1 - \frac{\Sigma_i \mid B_{ipred} - B_{iact} \mid}{N} \qquad (18)$$

where B_i stands for Basket i and N is the total number of shopping trips. This metric lies between 0 (no predictive capability) and 1 (perfect prediction).

As can be seen from Table 8, the proposed model has a hit rate of about 0.99 while the null models have hit rates of 0.85, 0.90, and 0.96, respectively.

A more detailed examination of the predictions shows that the independent probits model (null model 1) is poor at predicting all baskets. To illustrate why

this may be so, let us examine the case of the null basket. It may be expected that predicting the null basket is straightforward since 80% of the trips result in the null basket. However, recreating this null basket from a series of independent models does not account for the *joint* probability of no purchase in each category. Thus, though null model 1 performs very well at predicting no purchase/purchase for each category taken one at a time,[19] it cannot do so when the joint purchase is being predicted. This is because it ignores the complementarity and co-incidence linkages that drive multicategory incidence. The addition of complementarity (null model 2) and co-incidence (null model 3) both lead to an improvement in the Hit Rate with better predictions on all baskets. Of the two, complementarity seems to have lower impact than co-incidence in improving the Hit Rate. The combination of complementarity and co-incidence (proposed model) shows considerable improvement in predicting both single and joint category incidence (e.g., see baskets 4 and 11). A similar pattern can be seen in the predictions made by the proposed and null models using the holdout sample (Table 9).

Some researchers have argued that if multi-item choice models do not make better predictions than single-item choice models, the justification for developing these models is not clear (Russell et al. 1997). The results from the tests carried out above show that, in the context of our research, there are significant gains to be made from such models.

6. Managerial Implications

Our model provides many potential benefits to managers, especially in the retailing context. First, the examination of the correlation structure can enable retailers to identify traffic generating categories, i.e., categories which have a high amount of co-incidence with many other categories but few complementary effects. Second, retailers can devise better promotion policies for categories that are complementary. Third,

[19]There are very minor differences across the four models in terms of predicting single category purchase. The proposed model and null model 1 make slightly better predictions than either null model 2 or null model 3.

the pattern of correlations can help in reducing the entire set of categories into more manageable "chunks" for further analysis of brand choice, purchase quantities, and other phenomena of interest. Finally, the capability of generating household level estimates allows retailers to design direct marketing programs such as cross promotions (e.g., cross couponing) or higher across-the-board promotions to more price sensitive households. Manufacturers who have brands in complementary categories can also get similar insights from our model.

We demonstrate the usefulness of our model by assessing the effect of retailer pricing policies under different model assumptions. We assume that the retailer would like to quantify the effect of price promotions on profit. We further assume that the retailer drops price on each of the four categories by 10% as his or her promotional strategy. Since we have no knowledge of the exact margins in these categories, we look at a range of gross margins (20% to 60%). These margins encompass numbers reported in industry publications such as *Progressive Grocer*.[20] We assume that the total market consists of the 155 households that we have used in our calibration sample. We use the results from null model 1, null model 2 and the proposed model to estimate the profit impact (detailed in Table 10).

The results show that each model gives a different estimate of the final profit. Looking at the numbers in detail, we find that the addition of cross-effects (null

[20]Since the purpose of these analyses is to compare profits across models, the absolute numbers are not very important.

Table 10   Retail Profits

Gross Margin (%)	Null Model 1[1] Profit ($)	Null Model 2[2] Profit ($)	% Increase over Null Model 1	Proposed Model Profit ($)	% Increase over Null Model 1
20	235.33	243.74	4	259.92	10
30	460.68	480.98	4	507.85	10
40	695.05	723.33	4	766.76	10
50	921.43	966.58	5	1021.68	11
60	1119.81	1166.84	4	1232.89	10

[1]Independent probit models.

[2]Independent probit models with complementarity.

model 2) translates into an average increase in estimated profits of 4% relative to the profits estimated from the independent probits model (null model 1). Adding co-incidence to null model 2, i.e., using the proposed model, adds another 6% to the same base, i.e., the proposed model shows that profits may be increased by about 10% over that of null model 1. We should highlight that we are not optimizing profits, but simply uncovering the multicategory profitability of promotions. If a retailer misjudges the profit potential of a promotion, he/she may forego profitable opportunities. For example, a manager may use null model 1 to evaluate the profitability of a specific promotion and may decide that it was not profitable enough. However, if the manager were to use the proposed model, the incremental profit due to complementarity might tip the decision in favor of running the promotion.

7.   Conclusions and Future Research

In this research we propose a general approach to modeling the shopping basket. The composition of the basket is of significant interest to retailers and manufacturers. We also propose a general framework for classifying multicategory choice situations in terms of complementarity and co-incidence (Table 1) and we show how these problems may be modeled through the use of an appropriate model form. The shopping basket is one of many multicategory choice phenomena that are encountered by consumers. This modeling approach is likely to prove very useful in other applications such as database marketing and micromarketing. Finally, the multivariate probit model may also be applied in any marketing situation in which there could be more than one outcome simultaneously, e.g., it could be used to model the consideration set of consumers in the brand choice process.

Our results show that it is possible to separate the controllable drivers of multicategory choice from the uncontrollable ones. We also show that, in some cases, the addition of these controllable explanatory variables reduces the effect of the uncontrollable variables. In addition to separating these effects, we are also able to quantify the magnitude of these effects. In general, as

expected, complementarity effects are smaller in magnitude than the own effects. However, since the grocery retailing industry is highly competitive with low margins of around 5% (*Food Marketing Institute* 1992), even small demand (and profit) gains through cross pricing may be of great significance. The nature of the co-incidence pattern can provide insights into how the bigger problem of multi-category choice may be reduced to many smaller problems of interdependent pairs of items. This is likely to be of much value to retailers and manufacturers. The use of hierarchical models also allows us to capture heterogeneity at the individual household level—this may enable managers to formulate and implement micromarketing strategies. In terms of applicability of our results, we would like to note that our sample consists of average to heavier buying households across the four categories that we have chosen. Thus, these insights are likely to be limited to these kinds of households and may not be applicable to all households that shop at a particular store or stores.[21]

Our current and future work in this area is in four broad directions. In the first stream we plan to model other multicategory decisions such as brand choice and purchase quantity as well as build integrated models which link these decisions. In the second stream, we plan to develop a theoretical model of multicategory consumer choice under an axiomatic, utility maximization framework. We would then transform the theory's predictions into an empirically estimable model which would use methods such as the ones proposed in this paper. For an excellent example of this approach in the single category context, see Arora et al. (1998). In the third stream, we investigate the nature of household heterogeneity across categories in more detail and try to build models for micromarketing applications. Finally, we are also investigating the extension of these models into other substantive areas such

as database marketing and the choice of households' financial portfolios.[22]

Appendix 1.   Prior Distributions
1. The prior distribution of γ is a multivariate normal distribution MVN[η, C] where $\eta = 0$ and $C = \text{diag}(10^3)$.

2. The prior distribution of μ_0 is a multivariate normal MVN[θ_0, Θ_0] where $\theta_0 = 0$ and $\Theta_0 = \text{diag}(10^3)$.

3. The prior distribution of μ_1 is a multivariate normal MVN[θ_1, Θ_1] where $\theta_1 = 0$ and $\Theta_1 = \text{diag}(10^3)$.

4. The prior distribution of μ_2 is a multivariate normal MVN[θ_2, Θ_2] where $\theta_2 = 0$ and $\Theta_2 = \text{diag}(10^3)$.

5. The prior distribution of Λ_0^{-1} is Wishart W[ρ, $(\rho R_0)^{-1}$] where ρ = NCAT + 1 and $R_0 = \text{diag}(10^{-3})$. NCAT is the number of categories, i.e., in this application NCAT is 4.

6. The prior distribution of Λ_1^{-1} is Wishart W[ρ, $(\rho R_1)^{-1}$] where ρ = NCAT + 1 and $R_1 = \text{diag}(10^{-3})$.

7. The prior distribution of Λ_2^{-1} is Wishart W[ρ, $(\rho R_2)^{-1}$] where ρ = NCAT + 1 and $R_2 = \text{diag}(10^{-3})$.

8. The prior distribution for the free elements of the correlation matrix (Σ), vec*(Σ) = (σ_{12}, σ_{13}, ... $\sigma_{(J-1)J}$)', such that vec* (Σ) $\in A$, can be expressed as

$$\pi(\text{vec*}(\Sigma) \mid \Sigma_0, G_0) \propto \exp$$

$$\left\{ -\frac{1}{2} (\text{vec*}(\Sigma) - \text{vec*}(\Sigma_0))' G_0 (\text{vec*}(\Sigma) - \text{vec*}(\Sigma_0)) \right\},$$

where the region A is a subset of the region $[-1, 1]^{J(J-1)/2}$ that leads to a proper correlation matrix, Σ_0 is a $J \times J$ correlation matrix with all diagonal elements equal to 0 and G_0 is a $(J(J-1)/2) \times (J(J-1)/2)$ precision matrix. Note that Σ_0 and G_0 are the hyperparameters for the prior. We set $\Sigma_0 = 0$ and $G_0 = I$ where I is an identity matrix of dimensionality $(J(J-1)/2) \times (J(J-1)/2)$.

Appendix 2.   Description of Full Conditionals
1. The full conditional distribution for γ is multivariate normal, $N[\hat{\gamma}, E^{-1}]$, where $\hat{\gamma} = E^{-1} (C^{-1}\eta + \Sigma_{h=1}^{H} \Sigma_{t=1}^{n_h} Z_{ht}' \Sigma^{-1} \tilde{u}_{ht})$ and $E = C^{-1} + \Sigma_{h=1}^{H} \Sigma_{t=1}^{n_h} Z_{ht}' \Sigma^{-1} Z_{ht}$. The adjusted utilities, $\tilde{u}_{ht}$, are defined as $u_{ht} - \beta_{h0} - X_{h1t}\beta_{h1} - X_{h2t}\beta_{h2}$.

2. The full conditional distribution for β_{h0} is multivariate normal, $N[\hat{\beta}_0, B_0^{-1}]$, where $\hat{\beta}_0 = B_0^{-1} (\Lambda_0^{-1} D_h \mu_0 + \Sigma_{t=1}^{n_h} \Sigma^{-1} \hat{u}_{ht})$ and $B_0 = (\Lambda_0^{-1} + \Sigma_{t=1}^{n_h} \Sigma^{-1})$. The adjusted utilities, $\hat{u}_{ht}$, are defined as $u_{ht} - Z_{ht}\gamma - X_{h1t}\beta_{h1} - X_{h2t}\beta_{h2}$ for all h and t.

---

[21]Though this (restricted) applicability may be seen as a limitation, retailers are keenly focussed on the buying behavior of these "heavier" buyers. This is because these customers buy larger baskets on each visit and therefore have a higher impact on sales and profits *across* categories, e.g., Ukrop Supermarkets labels customers who spend about twice as much as the average customer during a store trip as its "Valued Customers" and provides them extra promotional incentives (*Food Marketing Institute* 1994).

3. The full conditional distribution for β_{h1} is multivariate normal, $N[\hat{\beta}_1, B_1^{-1}]$, where $\hat{\beta}_1 = B_1^{-1} (\Lambda_1^{-1} D_h \mu_1 + \Sigma_{t=1}^{n_h} X'_{h1t} \Sigma^{-1} \hat{u}_{ht})$ and $B_1^{-1} = (\Lambda_1^{-1} + \Sigma_{t=1}^{n_h} X'_{h1t} \Sigma^{-1} X_{h1t})$. The adjusted utilities, $\hat{u}_{ht}$, are defined as $u_{ht} - \beta_{h0} - X_{h2t} \beta_{h2} - Z_{ht} \gamma$ for all h and t

4. The full conditional distribution for β_{h2} is multivariate normal, $N[\hat{\beta}_2, B_2^{-1}]$, where $\hat{\beta}_2 = B_2^{-1} (\Lambda_2^{-1} D_h \mu_2 + \Sigma_{t=1}^{n_h} X'_{h2t} \Sigma^{-1} \hat{u}_{ht})$ and $B_2^{-1} = (\Lambda_2^{-1} + \Sigma_{t=1}^{n_h} X'_{h2t} \Sigma^{-1} X_{h2t})$. The adjusted utilities, $\hat{u}_{ht}$, are defined as $u_{ht} - \beta_{h0} - X_{h1t} \beta_{h1} - Z_{ht} \gamma$ for all h and t.

5. The full conditional distribution for μ_0 is multivariate normal, $N[\hat{\mu}_0, M_0^{-1}]$, where $\hat{\mu}_0 = M_0^{-1} (\Theta_0^{-1} \theta_0 + \Sigma_{h=1}^{H} D'_h \Lambda_0^{-1} \beta_{h0})$ and $M_0^{-1} = \Theta_0^{-1} + \Sigma_{h=1}^{H} D'_h \Lambda_0^{-1} D_h$.
The full conditional distributions for μ_1 and μ_2 are also multivariate normal and may be obtained in a similar manner.

6. The full conditional distribution for Λ_0^{-1} is Wishart and is given by $W\{(\rho R_0 + \Sigma_{h=1}^{H} (\beta_{h0} - D_h \mu_0)(\beta_{h0} - D_h \mu_0)')^{-1}, \rho + H\}$ where H is the number of households.
The full conditional distributions for Λ_1^{-1} and Λ_2^{-1} are also Wishart and may be obtained in a similar manner.

7. The full conditional distribution for the identified set of utilities on any observation is truncated multivariate normal, where the region of truncation is determined by the pattern of incidence across the categories. The truncated multivariate normal variate can be generated by using a mini-Gibbs sampler (Geweke 1991). This involves using J Gibbs sampling steps such that the jth component of u_{ht}, u_{hjt}, comes from the univariate truncated conditional normal density $p(u_{hjt} | \beta_{h0}^{(m)}, \beta_{h1}^{(m)}, \beta_{h2}^{(m)}, \gamma, \Sigma^{(m)}, y_{ht})$. Here the univariate normal distribution is truncated from below by 0 if a purchase was made in the category j, else it is truncated from above by 0.

8. The final step of the substitution sampler involves sampling the Σ correlation matrix. If the prior distribution for the freeelements of the correlation matrix Σ, vec* $(\Sigma) = (\sigma_{12}, \sigma_{13}, \ldots \sigma_{(J-1)J})'$, is $\pi(\text{vec*}(\Sigma) | \Sigma_0, G_0)$, then the full conditional distribution of Σ is proportional to $\{L(\Sigma | \beta_{h0}^{(m)}, \beta_{h1}^{(m)}, \beta_{h2}^{(m)}, \{\{u_{ht}\}\}, y)\} \tau(\text{vec*}(\Sigma) | \Sigma_0, G_0)$ where $L(.)$ is the conditional likelihood and is given as (ignoring the normalizing constant).

$$|\Sigma|^{-n/2} \exp\left\{ -\frac{1}{2} \sum_{i=1}^{n} (u_i - \beta_{0i} - x_{1i}\beta_{1i} - x_{2i}\beta_{2i} - z_i\gamma_i)' \right.$$

$$\left. \Sigma^{-1} (u_i - \beta_{0i} - x_{1i}\beta_{1i} - x_{2i}\beta_{2i} - z_i\gamma_i) \right\},$$

where vec* $(\Sigma) \in A$ as in Appendix A1.

Since this full conditional distribution is complex, we generate Σ using a Metropolis Hit-and-Run algorithm (Dey and Chen 1996, Chen and Schmeiser 1993). If $\Sigma^{(m)}$ is the current value of the correlation matrix, then in this step, a candidate matrix Σ^c is generated by specifying a random walk chain $\Sigma^c = \Sigma^{(m)} + H$, where $H = (h_{ij})$ is a matrix such that $E(h_{ij}) = 0$ and $h_{ii} = 0$. Let ξ be the smallest eigenvalue of $\Sigma^{(m)}$. Then the elements of the increment matrix H can be generated using the Hit-and-Run algorithm. This involves the following steps:

(a) generate a sequence of *i.i.d* standard normal deviates z_{12}, $z_{13}, \ldots, z_{(J-1),J}$, of length $J(J - 1)/2$,

(b) generate a deviate d from $N(0, \sigma_d^2)$ which is truncated to the interval $(-\xi/\sqrt{(2)}, \xi/\sqrt{(2)})$,

(c) formulate the elements

$$h_{ij} = \frac{dz_{ij}}{\left(\sum_{j=1}^{J-1} \sum_{l=j+1}^{J} z_{jl}^2 \right)^{(1/2)}}$$

for $i < j$, $h_{ii} = 0$, and $h_{ij} = h_{ji}$ for $i > j$.

Here σ_d^2 is a tuning constant that needs to be chosen such that candidates are not rejected disproportionately. If ξ^c is the smallest eigenvalue of the candidate matrix, then once a candidate is generated, it is accepted or rejected based on the following Metropolis-Hastings acceptance probability:

$$\min\left\{ \frac{L(\Sigma^c | \beta_{h0}^{(m)}, \beta_{h1}^{(m)}, \beta_{h2}^{(m)}, \{\{u_{ht}\}\}, y) p(\text{vec*}(\Sigma^c)) \left(\Phi\left(\frac{\xi^c}{\sqrt{2}\sigma_d}\right) - \Phi\left(\frac{-\xi^c}{\sqrt{2}\sigma_d}\right) \right)}{L(\Sigma | \beta_{h0}^{(m)}, \beta_{h1}^{(m)}, \beta_{h2}^{(m)}, \{\{u_{ht}\}\}, y) p(\text{vec*}(\Sigma)) \left(\Phi\left(\frac{\xi}{\sqrt{2}\sigma_d}\right) - \Phi\left(\frac{-\xi}{\sqrt{2}\sigma_d}\right) \right)}, 1 \right\}$$

where $\Phi(.)$ is the standard normal cumulative distribution function.

References

Ainslie, Andrew, Peter E. Rossi. 1998. Similarities in choice behavior across product categories. *Marketing Sci.* **17** 91–106.

Ansari, Asim, Maureen Morrin, Alok Gupta. 1996. Multi level models in marketing research. Working Paper. Graduate School of Business, Columbia University, New York.

Albert, J., S. Chib 1993. Bayesian analysis of binary and polychotomous response data. *J. Amer. Statis. Assoc.* **88** 669–679.

Allenby, Greg M., James L. Ginter. 1995. Using extremes to design products and segment markets. *J. Marketing Res.* **32** November 392–403.

Arora, Neeraj, Greg M. Allenby, James L. Ginter. 1998. A hierarchical Bayes model of primary and secondary demand. *Marketing Sci.* **17** 29–44.

Bell, David R., James M. Lattin. 1998. Grocery shopping behavior and consumer preference for retail price format: why 'large basket' shoppers prefer EDLP. *Marketing Sci.* **17** 66–88.

Berry, Jonathan. 1994. Database marketing. *Bus. Week* (September 5) 56–62.

Bodapati, Anand. 1996. The impact of out-of-store advertising on store sales. Unpublished Doctoral Dissertation. Stanford University, Stanford, CA.

Bucklin, Randolph E., Sunil Gupta. 1992. Brand choice, purchase incidence and segmentation: an integrated approach. *J. Marketing Res.* **29** 201–215.

Chen, Ming-Hui, B. W. Schmeiser. 1993. Performance of the Gibbs, hit-and-run, and metropolis samplers. *J. Comput. Graphical Statist.* **2** 251–272.

Chib, S., E. Greenberg. 1998. Analysis of multivariate probit models. *Biometrika* **85** 347–361.

——, E. Greenberg. 1995. Understanding the Metropolis-Hastings Algorithm. *Amer. Statistician* **49** 327–335.

Chintagunta, Pradeep, Sudeep Haldar. 1998. Investigating purchase timing behavior in two related product categories. *J. Marketing Res.* **35** 43–53.

Dey, Dipak K., Ming-Hui Chen. 1996. Bayesian analysis of correlated

binary data models. Technical Report 1996-02. Department of Statistics, University of Connecticut, Storrs, CT.

Donovan, Robert J., John R. Rossiter, Gilian Marcoolyn, Andrew Nesdale. 1994. Store atmospherics and purchasing behavior. *J. Retailing* **70** 283–294.

Fader, Peter, Leonard M. Lodish. 1990. A cross-category analysis of category structure and promotional activity for grocery products. *J. Marketing* **54** 52–65.

Food Marketing Institute. 1992. Alternative store formats: Competing in the nineties. Report 9-511, Washington, DC.

——. 1994. Exploring relationship marketing. Report published by the American Greetings Research Council, Cleveland, OH.

Geweke, J. 1991. Efficient simulation from the multivariate normal and student-*t* distributions subject to linear constraints. *Computing Science and Statistics: Proc. Twenty-Third Symposium on the Interface.* American Statistical Association, VA. 571–578.

Geyer, C. J. 1992. Practical Markov chain Monte Carlo. *Statist. Sci.* **7** 473–511.

Greene, William H. 1997. *Econometric Analysis.* 3rd Edition. Prentice Hall, Englewood Cliffs, NJ.

Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57** 97–109.

Hoch, Stephen J., Byung-Do Kim, Alan L. Montgomery, Peter E. Rossi. 1995. Determinants of store level price elasticities. *J. Marketing Res.* **32** 17–29.

Kahn, Barbara E., David C. Schmittlein. 1992. The relationship between purchases made on promotion and shopping trip behavior. *Marketing Letters* **68** 294–315.

Kim, Byung-Do, Kannan Srinivasan. 1995. Investigation of consumers' price sensitivity across multiple product categories: An empirical investigation. Working Paper. GSIA, Carnegie-Mellon University, Pittsburgh, PA.

Krishnamurthi, L., S. P. Raj. 1988. A model of brand choice and purchase quantity price sensitivities. *Marketing Sci.* **7** 1–20.

Manchanda, Puneet, Sunil Gupta. 1997. Complementarity in shopping baskets: investigating multi-category purchase incidence and brand choice. Working Paper. Graduate School of Business, University of Chicago, Chicago, IL.

McCulloch, Robert, Peter E. Rossi. 1994. An exact likelihood of the multinomial probit model. *J. Econometrics* **64** 207–240.

Menon, Satya, Barbara E. Kahn. 1995. The impact of context on variety seeking in product choices. *J. Consumer Res.* **22** 285–295.

Meyer, Robert J., Barbara E. Kahn. 1991. Probabilistic models of consumer choice behavior. Thomas S. Robertson, Harold H. Kassarjian, eds. *Handbook of Consumer Behavior.* Prentice-Hall, Englewood Cliffs, NJ 85–122.

Mogelonsky, Marcia. 1994. Please don't pick the cherries: how supermarketers use electronic price scanning to create loyal customers. *Marketing Tools* (September/October) 10–13.

Mulhern, Francis J., Robert P. Leone. 1991. Implicit price bundling of retail products: a multiproduct approach to maximizing store profitability. *J. Marketing* **55** 63–76.

Narasimhan, Chakravarthi, Scott A. Neslin, Subrata Sen. 1996. Promotional elasticities and category characteristics. *J. Marketing* **60** 17–30.

Park, C. Whan, Easwer Iyer, Daniel C. Smith. 1989. The effects of situational factors on in-store grocery behavior: the role of store environment and time available for shopping. *J. Consumer Res.* **15** 422–433.

Raftery, A. E., S. M. Lewis. 1995. How many iterations of the Gibbs sampler. J. M. Bernardo, J. O. Berger, A. P. Davis, A. F. M. Smith, eds. *Bayesian Statistics,* vol. 4. Oxford University Press, Oxford, U.K. 641–649.

Raju, Jagmohan S. 1992. The effect of price promotions on variability in product category sales. *Marketing Sci.* **11** 207–220.

Ratneshwar, S., Cornelia Pechman, Allan D. Shocker. 1996. Goal-derived categories and the antecedents of across-category consideration. *J. Consumer Res.* **23** 240–250.

Rossi, Peter E., Robert E. McCulloch, Greg M. Allenby. 1996. The value of purchase history data in target marketing. *Marketing Sci.* **15** 321–340.

Russell, Gary, David Bell, Anand Bodapati, Christina Brown, Jeongwen Chiang, Gary Gaeth, Sunil Gupta, Puneet Manchanda. 1997. Perspectives on multiple category choice. *Marketing Letters* **8** 297–305.

Swinyard, William R. 1993. The effects of mood, involvement and quality of store experience on shopping intentions. *J. Consumer Res.* **20** 271–280.

Terbreek, Glen. 1993. Why category management will fail. *Progressive Grocer* (September) 11–13.

Walters, Rockney G. 1991. Assessing the impact of retail price promotions on product substitution, complementary purchase, and interstore displacement. *J. Marketing* **55** 17–28.