Faculty Research

University of Michigan Business School

WORKING PAPER SERIES

Assesing Cross-Category Impact from Store-Level Scanner Data

Michel Wedel
University of Michigan Business School

Jie Zhang
University of Michigan Business School

Working Paper 03-003

Research Support • Ann Arbor, MI 48109-1234 • 763-9266

Assessing Cross-Category Impact from Store-Level Scanner Data

Michel Wedel[1]

and

Jie Zhang[2]

December 2002

Acknowledgement:

We are grateful to the editor and three anonymous reviewers for their many helpful suggestions and comments that have lead to substantial improvement of the paper.

[1] University of Michigan Business School, 701 Tappan Street, Ann Arbor, MI 48109, USA. (wedel@bus.umich.edu)

[2] University of Michigan Business School, 701 Tappan Street, Ann Arbor, MI 48109, USA. (jiejie@umich.edu)

Assessing Cross-Category Impact from Store-Level Scanner Data

ABSTRACT

We develop an approach to assessing cross-category effects of marketing mix variables from store level scanner data, which aims at providing manufacturers with insights to coordinate marketing activities across categories. We intended to provide a methodology to estimate all cross-category marketing effects, without imposing structural constraints on the coefficients. We allow all own and cross-effects within a category and across different categories to differ from brand to brand. As a solution to the problem of estimating a large number of parameters, we provide a Bayesian approach to pooling the effects to the category average, which tends to reduce problems of instability of the individual coefficients. We account for heterogeneity in baseline sales using an extended Dirichlet Process Prior-based method which accommodates the effects of demographics, and for longitudinal shifts in baseline sales through a semi-parametric Bayesian version of local polynomial regression. We analyze store level scanner data on two related product categories and illustrate the practical value of our approach for manufacturers' cross-category decisions.

Key Words: Local Polynomial Regression, Dirichlet Process Prior, Bayesian Pooling, Semi-Parametrics, Non-Parametrics.

INTRODUCTION

Improved understanding of cross-category competitive effects enables manufacturers to allocate marketing budgets across categories, to preempt or retaliate competition, and to enhance the effectiveness of cross-category branding, pricing and promotion strategies. The growing importance of cross-category managerial decision making has prompted a stream of academic research in recent years (*e.g.*, Walters 1991; Russell and Kamakura 1997; Ainslie and Rossi 1998; Chintagunta and Haldar 1998; Erdem 1998; Manchanda, Ansari, and Gupta 1999; Song and Chintagunta 2002). Although providing important insights into consumers' cross category behavior, the studies done to date do not directly enable manufacturers to answer simple questions such as: "How does the change in my brand's price in one category affect the sales of my brand in another category?" or, "What happens to the sales of my brand if a competitor's brand in a related category changes its regular price or runs a price promotion?" From a brand managers' perspective, insight into competitive and cannibalization effects for individual brands in different categories is critical for her decisions on the allocation of marketing effort across categories.

We intend to develop a model that assesses with- and between-category effects of marketing mix on sales *at the brand level*, and thus provides direct answers to the above questions. The model is calibrated on store level scanner data. There are three major challenges in assessing cross-category market response for individual brands from store-level scanner data. First, the models in question involve a large number of parameters. For two categories with K brands in each, there are $2K^2$ within-category own- and cross-effects and $2K^2$ between-category own- and cross-effects, for each marketing mix variable. Estimating such large number of parameters directly often yields unstable and counterintuitive estimates (Blattberg and Wisniewski 1988; Boatwright, McCulloch and Rossi 1999). Prior studies dealt with this problem by limiting the number of brands in each category (*e.g.*, Walters 1991), assuming the own- and cross-effects to be the same across brands (*e.g.*, Ainslie and Rossi 1998; Erdem and Winer 1999), assuming them to be proportional to market share (*e.g.*, Kopalle, Mela and Marsh 1999), or constraining the parameter estimates to have the correct sign (*e.g.*, Boatwright, McCulloch and Rossi 1999). Our primary intended contribution is to provide an approach to estimating all within and between category own- and cross-elasticities without imposing structural constraints on the parameters, assuming them to be homogeneous across stores. We use a Bayesian approach to shrink the estimates towards their category averages. Bayesian shrinkage approaches in which coefficients are pooled across *different equations* have been previously used (*e.g.* Blattberg and George 1991; Boatwright, McCulloch, and Rossi 1999). Although serving a different purpose, what is novel in our study is that we pool the cross-effects of marketing mix variables *within a single regression equation*, to reduce problems of instability of the individual coefficients and occurrence of counterintuitive signs. Our specification is based on the assumption that the own- and cross-effect coefficients within and between categories are draws from a distribution with common category-level means.

Second, it is important to account for heterogeneity in sales levels across stores. The marketing literature has extensively documented the importance of dealing with heterogeneity (Allenby and Rossi 1999). Approaches that account for heterogeneity by imposing a distribution of store-level parameters are superior to fixed effects approaches which we estimate a separate parameter for every store (cf. Kopalle, Mela and Marsh 1999). This is because fixed-effect estimation conditions on the observed sample and does not allow for inference on the population of stores, and it is also statistically less efficient. Bayesian hierarchical models and finite mixture models have become standard techniques to handle heterogeneity (Allenby and Rossi 1999), but their performance is sensitive to violations of assumptions on the heterogeneity distribution. Research has shown that general support exists neither for a discrete nor for a continuous representation (Edwards, Ansari and Currim 2001). We use a nonparametric Bayesian method based on the Dirichlet Process Prior (DPP) to alleviate the problems associated with the above two standard techniques (cf. Ferguson 1973; Escobar 1994). This approach is more versatile because it nests both of the above representations and adapts to the shape of the actual density empirically (Escobar 1994). In addition, we extend the standard DPP approach by including the effects of store demographics in the trade area on sales. See Hoch *et al.* (1995) for theory and prior research on the effects of such variables.

Third, failure to account for nonlinear trends in category sales may render biased estimates of cross-category effects of marketing mix. Some previous studies on store level sales data have used weekly dummy variables to capture such nonlinear trends in sales at the category level (*e.g.*, Wittink, Addona, Hawkes and Porter 1988; Foekens, Leeflang and Wittink 1994; Kopalle, Mela and Marsh 1999). This approach, however, greatly adds to the number of parameters estimated in the models and may overcorrect if there is a smooth seasonal trend. Abraham and Lodish (1987, 1993) use pre-smoothing of the sales time-series to remove time trends before estimating market response models. While this method is more parsimonious, it may result in over-smoothing of the data and in removing part of the structural effects of the marketing mix variables of interest. To avoid these problems, we develop a semi-parametric representation of category level sales over time, using a Bayesian version of local polynomial regression (Fan and Gijbels 1997). The power and flexibility of semi-parametric models have been demonstrated by, for example, Van Heerde, Leeflang and Wittink (2001) and Kalyanam and Shively (1998) in the marketing literature.

The paper is organized as follows. In the next section, we offer a brief review of the relevant literature, compare our study with extant ones, and elaborate its contributions. Next, we present the cross-category model and its estimation. Then, we provide a detailed conceptual framework of plausible routes through which cross-category effects may take place. We illustrate our approach in a case study using weekly store-level sales data of refrigerated and frozen orange juice from a single retail chain, demonstrate its practical value to manufacturers' cross-category price and promotion decisions, and conclude with discussions of general managerial implications and limitations of the study.

RELEVANT CROSS-CATEGORY RESEARCH LITERATURE

The cross-category research literature has predominantly focused on understanding dependencies in individual consumers' purchase behavior across categories. These studies have focused on examining the *correlations* of consumers' brand preferences (Russell and Kamakura 1997; Erdem 1998; Erdem and Winer 1999), responses to marketing mix variables (Ainslie and Rossi 1998), state dependence (Seetharaman, Ainslie, and Chintagunta 1999), and purchase timing or incidence probability (Chintagunta and Haldar 1998). The consumer behavior studied ranges from category purchase incidence/timing (Manchanda *et al.* 1999; Russell and Petersen 2001; Ansari and Iyengar 2002), brand choice (Ainslie and Rossi 1998; Erdem 1998; *etc.*), to purchase quantity (Russell and Kamakura 1997). Household panel data have been the most common type of data used in the cross-category research literature. In the models calibrated on panel data, it is common to assume the cross-category effects of marketing variables the same for different brands in a category (*e.g.*, Ainslie and Rossi 1998; Erdem and Winer 1999; Russell and Petersen 2000).

In order to assess the *overall* cross-category *sales* effects, we argue that one needs to utilize *store-level* data. Although household panel data have provided rich insights into cross-category dependencies in individual consumers' purchase behaviors, they do not reflect cross-category effects caused by, amongst others, changes in store traffic and retail chain level category expansion due to acquisition of new users. Store level data do reflect these effects, and thus are more preferable if one wants to assess the impact of a brand's marketing mix on sales of all brands in another category, like in our study[3]. Only a few studies on cross-category effects to date, however, have used such store-level sales data (*e.g.*, Walters 1988 and 1991; Song and Chintagunta 2002). Further methodological developments in analyzing store level cross category data are thus needed, given its widespread use for marketing decisions in the industry (Leeflang and Wittink 2000).

Studies on cross-category effects using store level data all face the problem of how to estimate a large number of parameters. Walters (1991) investigates the impact of price promotions on sales of substitutes and complements within and across stores, but includes only a very small number of stores and brands[4]. Song and Chintagunta (2002) develop a random coefficients logit model that is calibrated on store-level data to investigate the dependence of marketing mix effects across several categories, arising from store choice, category incidence, and brand choice. Their model assumes that marketing effects across brands are homogeneous, and cross-category effects arise only from store choice and category purchase incidence

[3] In addition, Gupta *et al.* (1996) show that panel households may not be representative of the population of consumers and that analysis of household purchase records may provide price elasticity estimates quite different than those from store data.

decisions. The random effects specification does lead to different cross-effects among brands, but their structure is constrained by the nature of the formulation which is based on an assumption of independence of the random variation of different marketing mix effect parameters.

A practical reason for assuming all cross-effects to be homogeneous among brands in a category is that it renders the models more parsimonious and thus more suited for the analysis of a larger number of categories (cf. Song and Chintagunta 2002). But, this assumption is restrictive in terms of the implied managerial actions, and is likely to yield weak cross-category effects in empirical analyses as a result of averaging the individual brand effects. The empirical estimation of all brand-specific effects is a nontrivial task, especially in cross-category studies in which the number of parameters increases drastically if one is interested in all own- and cross-effects within and between categories. Hoch *et al.* (1995) and Montgomery (1997) have proposed techniques to estimate all own- and cross-effects in a single category. But, they do not look into effects across categories, where that problem is more salient. An approach to assessing brand-specific within- and cross-category effects is especially useful for brand managers who are interested in coordinating decisions across related categories, and thus demand precise knowledge on what brands are affecting and being affected by their marketing activities and to what extent.

The core contribution of our study is to provide a methodology of estimating all own and cross-effects in a cross-category context without imposing structural constraints on them, while also accommodating store heterogeneity and longitudinal shifts in the baseline sales levels. While we do not intend to analyze a large number of categories, we set out to provide detailed

[4] It includes two stores, three brands in one category and two in another.

insight into all own- and cross- effects of marketing variables at the brand level, within and across categories. In terms of the specific marketing mix variables investigated, we focus on regular price and price promotion, although the approach can be well applied to other variables. We intend to provide a methodology that presents a decision-making aid to brand managers not yet offered by available methods.

THE CROSS-CATEGORY MODEL AND ITS ESTIMATION

We develop a cross-category model for store level sales data, starting from the popular log-log demand model. In the marketing literature double- and single-log demand models have been frequently used (*e.g.*, Wittink, Addona, Hawkes and Porter 1988; Foekens, Leeflang and Wittink 1994; Hoch et al. 1995; Montgomery 1997; Boatwright, McCulloch, and Rossi 1999), in spite of the fact that they do not lend to interior solutions to a retailer's category profit optimization decisions[5]. Note that the substantive focus of our study is on *manufacturers*' coordination of marketing activities across related categories.

We are interested in estimating all own- and cross-effects within and between the categories *at the brand level*, at the same time accommodating store heterogeneity, time trends in category sales and coincidence in sales due to unobserved factors. Since we do not want the latter two to confound the estimates of own and cross price effects, we use nonparametric and semi-parametric approaches, respectively, to represent them. Our model is formulated and estimated in the Bayesian framework. In most cases, our priors are standard conjugate non-

---

[5] This arises if cross effects of a marketing mix variable, say price, are included and their magnitude is substantial, and if the model does not accommodate store switching or retail competition. In that case, the total profit from a category based on the model can go to infinity by setting the price of one or more brands to infinity, because the cross effects will drive sales of the other brands to infinity.

informative priors. We will provide a motivation for the few informative priors that we choose to use.

Store Level Sales Model

Let $r = 1, \ldots, R$ indicate stores , $c = 1,\ldots,C$ indicate categories, $i = 1, \ldots, I_c$ indicate brands in category c, $t = 1, \ldots, T$ indicate weeks, and $k = 1, \ldots, K$ indicate marketing mix variables. (In our empirical application, we have $C = 2$ categories and $K = 2$ marketing mix variables, regular price and price discount). The model describes the log-sales $q_{r,t,c} = [q_{r,t,c,i}]$, as a function of the brand's own marketing mix, as well as of the marketing mix of the other brands in the same category and of those in the other category, contained in the partitioned row vector $x_{r,t,k} = [x_{r,t,k,c}]$, with $x_{r,t,k,c} = [x_{r,t,k,c,i}]$. We have:

$$\ln(q_{r,t,c}) = \alpha_{r,c} + \ln(x_{r,t})B_c + \psi_c(t) + \varepsilon_{r,t,c}, \tag{1}$$

where $\alpha_{r,c} = [\alpha_{r,c,i}]$ are store-specific brand constants, representing baseline sales, $\mathrm{B}_c = [\mathrm{B}_{c,d}]$ for $d = 1,\ldots,C$ indicates categories, and $\mathrm{B}_{c,d} = [[[\beta_{c,d,i,j,k}]]]$ is a matrix of category-specific coefficients. The double-log specification in (1) renders these coefficients directly interpretable as elasticities. Details of the model, its parameters and prior specifications are provided in the following sections.

Nonparametric Heterogeneity in Baseline Sales

We explain store-level variation in sales by store demographics and represent the unobserved heterogeneity through a nonparametric distribution for the brand constants. We use the Dirichlet Process Prior (DPP) based nonparametric approach developed by Ferguson (1973), Escobar (1994), and Escobar and West (1995)[6]. It is assumed that the constants $\alpha_{c,r}$, capturing the baseline sales, are distributed according to some unknown distribution function F. Rather than assuming a parametric form for F, such as the normal, we put a prior distribution on the class of possible distribution functions through the Dirichlet Process Prior, $DP(\rho, F_0)$, which is parameterized by a positive parameter ρ and prior distribution F_0. The parameter ρ determines the concentration of the prior around F_0 and has a prior Gamma distribution. We extend the standard DPP approach by specifying the prior to be a function of store-level demographic variables, $z_r = [z_{r,p}]$. Thus we have the following hierarchical setup:

$$
\begin{aligned}
&\alpha_{r,c,i} \sim F \\
&F \sim DP(\rho, F_0) \\
&\rho \sim G(r, q) \\
&F_0 = N(z_r' a_{c,i}, V_c)
\end{aligned}
\quad , \tag{3}
$$

where we use a Normal prior for F_0. Here, $a_{c,i} = [a_{c,i,p}]$, and V_c is a full $(I \times I)$ matrix. A conjugate inverse-Wishart prior $IW(10, I)$ is used for V_c, a multivariate Normal prior $N(0, 10^4 \times I)$ for $a_{c,i}$, and a $G(1,1)$ for ρ. Under these assumptions and priors, $\alpha_{r,c,i}$ come from a Dirichlet mixture of Normal distributions, with $L \leq R$ components, which provides a nonparametric approximation to distributions of arbitrary form. We use a somewhat informative prior for ρ with $E[\rho] = 1$ and 2 prior degrees of freedom[7], which assumes around five components in the DP mixture *a priori* (Escobar 1994, Table 1), because we intend to remove store level heterogeneity from the estimates of the own- and cross-elasticities.

[6] The DPP framework has been applied in marketing in working papers by Ansari and Iyengar (2002), Ansari and Mela (2002) and Kim, Menzefricke and Feinberg (2002).

Category Sales Variation over Time

We capture time-dependent variation in category sales through a smooth function, $\psi_c(t)$[8]. Previous researchers have used lagged-sales specifications (*e.g.*, Hoch *et al.* 1995), dummy variables (Foekens, Leeflang and Wittink 1994; Van Heerde, Leeflang and Wittink 2001), or pre-smoothing of the sales series (Abraham and Lodish 1987, 1993), to capture time trends in baseline sales. We prefer to build a more flexible time-series smoother into our model that is capable of capturing category-specific seasonal variation, category expansion and other unaccounted variation in sales. This is done through a Hierarchical Bayes extension of local polynomial regression (Fan and Gijbels 1996). We assume that the time-sales function can be locally approximated in a neighborhood $(-h, h)$ of time point t by a Taylor series expansion:

$$\psi_c(t) \approx \sum_{p=0}^{P-1} \int_{t-h}^{t+h} (z-t)^p \Omega_c dz \,. \tag{4}$$

This describes time variation locally by *P-1*[th] order local polynomials in the neighborhood *h*, where the coefficients of the local polynomial change stochastically over time according to a *P*-variate Wiener process with drift: $\Omega_c(s) = \omega_c + \Delta' B_c(s)$, where $B_c(s)$ is a standard *P*-variate Brownian motion, ω_c is a $(P \times 1)$ vector of means, and $V_\omega = \Delta'\Delta$ a $(P \times P)$ covariance matrix. $\Omega_c(dz)$ is the change in the process in an interval dz.

This development of the Hierarchical Bayes local polynomial regression is analogous to that of the stochastic spline regression applied by Kalyanam and Shively (1998), except that we use Bayesian shrinkage of all local polynomial regression coefficients rather than of the slope of the spline function. Fan and Gijbels (1996) show local polynomial regression to offer several

---

[7] Given 82 stores in our data, the influence of a prior with 2 degrees of freedom is mild.

[8] We note that this smooth function may not only capture seasonality and other time related aspects, but also the effects of variables not accounted for in our model.

advantages for semi-parametric analysis, including good efficiency properties and the absence of boundary effects. Equation (3) implies a uniform Kernel that results in an under-smoothed form with negligible bias, which is a desirable feature since we want to remove category time variation from sales. We use uninformative prior settings: $IW(10, I)$ for V_{ω}, and $N(0, 10^5 \times I)$ for ω_c.

Pooling of Within and Between Category Own- and Cross-Effects

We employ Bayesian shrinkage to estimate the own- and cross-elasticities contained in B_c. We prefer to use Bayesian shrinkage and informative priors over the use of "hard" constraints or placing zero prior mass on ranges of the parameter support, although such constraints, as an alternative approach, have been shown to work well for market response models for a single category (*e.g.*, Boatwright, McCulloch, and Rossi 1999). As will be elaborated later, cross-category effects take place in a more complex manner and in most cases their directions are difficult to predict *a priori*. Therefore, it is not desirable to use "hard" constraints or placing zero prior mass on part of the parameter space in models of cross-category effects. Our approach also helps stabilize the individual estimates by assuming that they are draws from distributions with common category level means, rather than treating the individual within and between category elasticities as fixed parameters. Specifically, we pool the individual own- and cross-effects towards their category means, assuming them to be homogeneous across stores, as follows:

(5) $$vec(B_{c,k,d}) \sim N(vec(M_{c,k,d}), w_{c,k} \cdot I),$$

where

$$M_{c,k,d} = \begin{cases} b_{c,k}^{(own)} \cdot I + b_{c,k}^{(cross)} \cdot (\iota\iota' - I), & c = d \\ b_{c,k}^{(other)} \cdot \iota\iota', & c \neq d \end{cases},$$

with I identity matrix of appropriate dimensionality, ι a column vector of ones and $b_{c,k}^{(own)}$, $b_{c,k}^{(cross)}$, $b_{c,k}^{(other)}$ and $w_{c,k}$ scalars.

We use informative, theory-based priors for the mean within-category regular price coefficients: $\left(b_{c,1}^{(own)}, b_{c,1}^{(cross)}\right)' \sim N\left((-2.75, 1.00)', diag\{10^{-2}, 10^{-2}\}\right)$, and prior variance $w_{c,k} \sim IG(4, 5)$, with 8 prior degrees of freedom and expectation of 0.8[9]. These priors are based on other studies analyzing the same product categories as in our empirical application, two orange juice categories (Hoch *et al.* 1995; Montgomery 1997; Song and Chintagunta 2002), which provide precision weighted mean regular price own- and cross-elasticities of -2.75, and 0.95, and a pooled variance of 0.8[10]. We have no previous results on price discount effects and cross-category elasticities. Because we expect the price discount elasticities to be somewhat smaller in magnitude than the regular price effects, we specify the following prior for them: $\left(b_{c,2}^{(own)}, b_{c,2}^{(cross)}\right)' \sim N\left((2.00, -0.75)', diag\{10^{-2}, 10^{-2}\}\right)$. Since the categories we analyze in the empirical application are likely to be substitutes, we specify a weakly informative prior for the cross category effects: $\left(b_{c,1}^{(other)}, b_{c,2}^{(other)}\right)' \sim N\left((0.5, -0.5)', diag\{10^{-4}, 10^{-4}\}\right)$. It is important to note that the prior specification does not imply symmetry of the individual cross elasticities, because the priors are formulated for the hyper-parameters, *i.e.*, the category-level elasticities in $M_{c,k,d}$, which does not enforce symmetry in the individual elasticities in $\mathrm{B}_{c,k,d}$. The hyper-

---

9 In the empirical application we found that the posterior distributions of the brand elasticities to the specification of the prior is very mild, due to the large quantity of data, but that the category elasticities are more sensitive to their specification, which is precisely our purpose.

parameters $b_c^{(own)}$, $b_c^{(cross)}$, and $b_c^{(other)}$ can be interpreted as pooled category-level own-, cross- and cross-category marketing mix elasticities and are comparable to the pooled estimates that are obtained in models assuming the same effects of marketing variables for all brands in a category.

Unobserved Correlation of Sales Patterns

Whereas the standard log-log demand model provides a series of independent regression equations for the different brands, our approach is based on multivariate regression (cf. Zellner 1972) with the assumption that the errors in the different brand sales equations may be correlated. Those correlations reflect associations in sales patterns across the categories unexplained by the other factors in the model. Thus, we postulate a multivariate normal distribution for the error terms across the sales equations for the two categories:

(5) $$\varepsilon_{r,t} \sim N(0,\Sigma),$$

where $\varepsilon_{r,t} = \left[\varepsilon'_{r,t,1},\ldots,\varepsilon'_{r,t,C}\right]'$ and $\varepsilon_{r,t,c} = [\varepsilon_{r,t,c,i}]$. A $IW(10,I)$ prior is used for Σ.

Estimation

To estimate the parameters, we use a MCMC algorithm in which the parameters are drawn recursively from their full conditional posterior distributions[11]. All the full conditionals take standard forms because of the use of conjugate priors. We use 15,000 draws in all analyses, a burn in of 5,000, and thin the target draws by 1 in 5. We monitor convergence of the Markov

---

[10] Hoch et al. (1995) do not report cross elasticities.

[11] We conducted several analyses on synthetic data to investigate the performance of our model and algorithm to a) represent smooth category specific sales time functions, b) provide a nonparametric approximation to the distributions of store-specific brand constants, and c) pool the own- and cross-effects of price and promotions within and between categories. The satisfactory performance on synthetic data enhances our confidence of its application to store level scanner data.

Chains using plots of the draws and predictive tests to investigate the performance of alternative models.

ANALYZING CROSS-CATEGORY EFFECTS

In our model, the between-category effects may take on a variety of directions due to various mechanisms through which a brand's marketing mix variable(s) may affect the sales of the brands in another category. Since the categories may act as complements or as substitutes, some of these effects may work in opposite directions so that the direction of a particular cross-category effect becomes an empirical issue. Although our model as detailed above is not constructed to *disentangle* them[12], it is important to *categorize* the plausible routes through which cross-category dependencies may take place, which helps better understand and utilize the patterns of cross-category effects revealed in empirical application of the model. Based on existing literature and our own postulates, we identify the following plausible mechanisms.

1. Category complementarity or substitutability: An increase in the attractiveness of a brand increases (decreases) the attractiveness of its complementary (substitute) category, thus potentially increasing (decreasing) sales of any brand, own or other, in the other category. This effect has been documented by Walters (1988, 1991) for store-level sales, and Chintagunta and Haldar (1998) for purchase timing.
2. Brand spill-over effect: The attractiveness of a brand in one category may increase the attractiveness of a product carrying the same brand name in another category. This effect has been well-documented (*e.g.*, Sullivan 1990; Montgomery and Wernerfelt 1992; Erdem 1998; Erdem and Winer 1999), and is the essence of umbrella branding.

[12] Song and Chintagunta (2002) investigate cross-category effects due to store switching and category purchase incidence separately.

3. Category spill-over effect: An increase in the attractiveness of one category may increase consumers' interest, and thus sales, of all brands in another category. We postulate this effect as an extension of the brand spill-over effect at the category level. Note that the difference between this effect and 1. is that it occurs for substitute and complementary products in the same direction. Chintagunta and Haldar (1998) have suggested this mechanism when trying to explain the positive cross-effect between liquid and powder laundry detergent.
4. Store traffic effect: An increase in a brand's attractiveness, such as through price promotion or feature advertisement, may draw shoppers from other stores and thus increase sales of the brand, as well as of products that are complements, substitutes, or unrelated. This effect has been described by Walters (1991), Bell and Lattin (1998), Chintagunta and Haldar (1998), Russell and Petersen (2000), and Song and Chintagunta (2002).
5. Budget constraint effect: The price a consumer pays for a brand in one category may affect the budget available for other categories and thus influence their sales. For example, a price promotion may increase the amount of money a consumer allocates to other categories and thus drive up sales of another category, while the reverse may happen to a raise in the regular price. This effect has been suggested by Chintagunta and Haldar (1998), Bell and Chiang (2002), and Song and Chintagunta (2002).

We present the expected signs of the above effects for price promotions in Table 1; regular price effects are opposite[13]. The table shows that cross-category dependence may occur

[13] While most previous research has aimed at understanding cross-category dependence from the perspective of consumer decision making, our categorization focuses on the resulting impact on a brand's sales. These perspectives

via multiple routes, and that some of the effects may work in opposite directions depending on the nature of the brands and categories. Therefore, to support manufacturers' cross-category marketing decisions, it is crucial to obtain estimates of cross-category effects *for each brand* separately. The framework in Table 1 may assist in interpreting the brand-specific parameter estimates.

A CASE STUDY FOR ORANGE JUICE

Data Description

We use store level scanner data from the Dominick's Fine Food chain based in the Greater Chicago area.[14]. Two product categories are used to calibrate our model: frozen orange juice concentrate, and refrigerated orange juice. These categories are treated as separate categories by retailers, such as Dominick's, as well as their manufacturers[15]. The data include weekly sales, regular prices and price discounts for each brand in the two categories in 82 stores over 104 weeks from January 1991 to December 1992. The top five brands of each category are used for calibrating the model: Citrus Hill, Florida Gold, Minute Maid, Tropicana, and Dominick's. These brands accounted for 100% of total frozen orange juice sales and 94.4% of total refrigerated orange juice sales in the stores and time period in our data. Using data of these two categories enables us to discuss cross-category coordination for all major manufacturers involved in the categories. In terms of positioning, Tropicana and Minute Maid are top tier

---

can be viewed as two sides of the same token. For example, consumers' learning experiences across categories can contribute to the brand spill-over effect and consumers' perception of products may lead to substitutability or complementarity. See Russell *et al.* (1999) for a review of cross-category dependence in brand choice behavior from the perspective of consumer decision making.

[14] We are grateful to The Kilts Center for Marketing of the University of Chicago Marketing Department (http://gsbwww.uchicago.edu/kilts/research/db/dominicks/dataset.shtml) for making the data available for academic usage.

[15] The extent to which these two categories are distinct in consumers' minds is a matter of degree, but our model estimation results confirm that they are indeed being perceived as distinct.

national brands and Citrus Hill and FloridaGold are second tier national brands in the refrigerated orange juice category, while Minute Maid is the top tier national brand and the other three national brands belong to a lower price tier in the frozen orange juice category. Dominick's is the private label brand. We also use six demographic variables of the trading area where a store is located to capture possible store-level sales variation due to demographics. These variables are: age (measured as percentage of population over age 60), education (measured as percentage of college graduates), income (measured as logarithm of median income), average household size (hhsize), percentage of detached houses (sinhouse), and population density of the trading area (in squared miles per capita). We use 52 weeks for model estimation and 52 weeks for holdout validation. Table 2 shows summary statistics of sales, regular prices and price discounts.

[INSERT TABLE 2 HERE]

Model Estimation Results

Figure 1 displays the category specific *smooth log-sales plots.* The plot reveals several spikes attributable to special occasions. In particular, sales spikes appear to occur around the Easter, Memorial Day, Labor Day, and Christmas holidays. Figure 2 shows the *density plots* for the brand-specific store constants in the two categories. The posterior means of the Dirichlet mixture parameters are $\rho = 76.44$ (SE = 24.06) and $\rho = 93.22$ (SE = 20.25), for the frozen and refrigerated juice categories, respectively. These parameters indicate that the distribution of the store constants is non-normal, which is corroborated by the plots in Figure 2. For a number of brands in both categories, the store constants exhibit bimodal distributions, while for the others the distribution of store constants is skewed, either to the right or to the left. These figures illustrate the advantage of using semi-parametric and nonparametric approaches to capture category specific variation in sales over time and heterogeneity in store level sales.

[INSERT FIGURES 1 AND 2 HERE]

Table 3 displays *demographic effects* on store sales. For frozen orange juice, only education significantly affects sales. Sales of all frozen juice brands are higher in trading areas with higher percentage of college graduates. This is also the case for refrigerated juice, but the effect is only significant for Tropicana and Dominick's. For refrigerated juice, sales of FloridaGold and Tropicana are significantly higher in trade areas with a higher percentage of people over 60 years old. Sales of the store brand Dominick's are significantly lower in trade areas with higher income. It appears that larger household size is associated with higher sales for all brands except Minute Maid. Finally, lower density trade areas are associated with lower sales levels of Tropicana and Dominick's.

[INSERT TABLE 3 HERE]

Table 4 shows the posterior means of the *pooled within and between category own- and cross-effects*. In both categories, all pooled within-category effects are significant and of expected signs. Compared to frozen juice, the regular price effects are smaller and the price discount effects larger for refrigerated juice, which may indicate that changes in regular price have more impact on sales for frozen orange juice, while price promotions are more effective for refrigerated orange juice, in general. The pooled cross-category effect parameters are small and insignificant, yet their precision (the inverse variance) is also low, which indicates that the individual cross-category effects may vary substantially across brands. This calls for a closer look at the individual brand level estimates.

[INSERT TABLE 4 ABOUT HERE]

Table 5 presents the posterior means of the *individual brand within-category effects*. The table reveals that the pattern of own and cross effects within each of the two categories is consistent with theory. The vast majority (85%) of the within-category effects are significant, and all but two of the significant effects are of expected signs. These two estimates (regular price effect of Citrus Hill on Tropicana and of Dominick's on Tropicana, both for refrigerated juice) are small in magnitude and could even be attributed to chance. Since the within category parameter estimates are intuitive and not the primary focus of our study, we will not discuss them in detail. Nevertheless, it is worth pointing out that the magnitude of regular price effects is larger than that of price discount effects in general. While most of the elasticities are of the same order of magnitude as those reported in previous studies on the orange juice category (*e.g.*, Hoch *et al.*, 1995; Song and Chintagunta 2002), we observe a few relatively large ones, in particular the large own regular price effect of Minute Maid, which may indicate that consumers are very sensitive to Minute Maid's regular price changes. The large parameter estimate could also be a result of separating regular price elasticities and price discount elasticities, which was not done in the above cited studies. Another point worth noting is that for both regular price and price discount, cross-effects tend to distribute more evenly across brand pairs for frozen juice, while revealing a more localized pattern among certain brand pairs for refrigerated juice, indicating a stronger hierarchy of brands in the latter category.

[INSERT TABLE 5 ABOUT HERE]

Table 6 presents the posterior means of the *individual brand cross-category effects*. Although the pooled cross-category estimates are insignificant, a substantial number of the individual cross-category effects are significant. They are in most cases smaller than the within category effects, which indicates that consumers tend to see these two categories as distinct. In addition, the signs of both types of price effects are positive as well as negative, which is consistent with our conjecture that the mechanisms driving these effects are more complex than mere substitutability of the two categories (see Table 1).

We first focus on the significant regular price effects. The framework in Table 1 (note that the signs of the regular price effects are exactly the opposite of the price promotion effects presented there) suggests that there should be more positive cross-regular price effects than own-regular price effects on a proportional basis, because a raise in a brand's regular price has a negative brand spill-over effect on own brand while positive brand spill-over effects on other brands in another category. This prediction is supported by our results in both categories. The positive cross-category regular price effects in Table 6 are mainly a result of substitutability of the two orange juice categories, plus an additional positive brand spill-over effect for the cross-effects. There are several negative cross-category regular price effects, which may indicate that, besides category substitutability, other effects such as brand spill-over, category spill-over, store traffic change, and budget constraints play a substantial role. Interestingly, the largest negative cross-regular price effects appear to occur to the highest priced brands in both categories, Minute Maid and Tropicana, which implies that these two brands are more likely to influence consumers' store choice, overall shopping budget, and create category spill-over. In terms of the specific brand pairs with negative cross-regular price effects, a general pattern is that a higher-priced brand's regular price reduction (increase) benefits (harms) a lower-priced brand in another category (*e.g.*, refrigerated Tropicana on frozen Dominick's, refrigerated Tropicana on frozen Citrus Hill, frozen Minute Maid and Tropicana on refrigerated Dominick's).

Next, we focus on the significant cross-category price discount effects. Table 1 suggests that there should be more negative cross-price promotion effects than own-price cross-category promotion effects on a proportional basis, because a price discount has a positive brand spill-over effect on own brand but negative brand spill-over effects on other brands in the other category. This prediction is supported in particular by the parameter estimates for frozen juice. The positive own- and cross-price discount effects in Table 6 suggest that factors other than category substitutability play a role in determining the cross-category effects between brands. Interestingly, FloridaGold always benefits from its promotion in the other category, while frozen Minute Maid and refrigerated Domnick's lose sales if promoted in the other category. This might be an indication that consumers of FloridaGold are more sensitive to umbrella branding, while consumers of the latter two brands see the two types of juice as closer substitutes. In addition, Dominick's refrigerated juice sales seem to benefit when other brands of frozen juice are on promotion, which suggests a larger store traffic effect for the store brand. We examine the implications of these different patterns on a manufacturer's cross-category decisions in the next section.

[INSERT TABLE 6 ABOUT HERE]

Table 7 shows the correlations between the residual log-sales, reflecting *the impact of unobserved factors on sales patterns across categories*. All brands, except for the store brand Dominick's, show a significant and positive correlation of their residuals across the two categories. This suggests that unobserved factors, such as brand image, affect sales of the same brand in both categories in the same direction, causing an additional brand-spill over effect.

[INSERT TABLE 7 ABOUT HERE]

We use the second year of data for m*odel comparisons* and examine *holdout forecast performance*. Log-sales in the holdout period of 52 weeks are predicted with a correlation of 0.805 (RMSE = 11.3% of mean log-sales) for frozen juice, and a correlation of 0.736 (RMSE =

11.1% of mean log-sales) for refrigerated juice. We consider the forecasting performance of the model quite satisfactory. To demonstrate the advantage of using the nonparametric Dirichlet Process Prior (DPP), we re-estimate the model with a Normal prior as a benchmark. In this model, the log-sales in the holdout period are predicted with a correlation of 0.785 (RMSE = 11.9%) for frozen juice, and of 0.702 (RMSE = 12.0%) for refrigerated juice, both of which are lower than those obtained with the model that includes the DPP heterogeneity representation. This indicates that the use of a nonparametric distribution for the store constants results in improved predictive validity relative to the standard method of normal distributions.

To examine the importance of estimating the parameters for each brand individually, we predict the 52 weeks holdout data using a model with only the pooled within- and between-category effects, omitting the individual own- and cross-elasticities, and retaining the remainder of the model structure, *i.e.*, the nonparametric store constants and semi-parametric time functions. The holdout correlations drop to 0.313 (RMSE = 65.7%) for frozen juice, and 0.437 (RMSE = 71.8%) for refrigerated juice, respectively. This drastic drop in the holdout correlations demonstrates that including the brand-specific effects not only provides incremental substantive insights, but also enhances predictive performance substantially.

Implications for Manufacturers' Cross-Category Decisions

Manufacturers are increasingly interested in coordinating marketing activities across related product categories. We illustrate how they can utilize the results of our model to design cross-category pricing and promotion strategies. There are four national brand manufacturers represented in our data which all produce refrigerated and frozen orange juice, making these two categories an excellent example to demonstrate the practical value of our methodology for manufacturers' cross-category decisions. Fundamentally, such decisions can only be made by considering the impact of a brand's price change or promotion on its own and competitors brands in both categories, *i.e.*, looking at all within- and between-category own- and cross-effects for each brand as presented in Tables 5 and 6.

For the manufacturer of Citrus Hill, it is more beneficial to reduce regular price of its frozen juice and run price promotions on its refrigerated juice. This is because a regular price reduction of frozen juice not only increases own brand sales and reduces the sales of other brands' within that category, but also increases own brand sales and hurts competitors' sales in the refrigerated juice category. On the contrary, a regular price reduction of refrigerated juice results in cross-category cannibalization, since it entices more Citrus Hill frozen juice consumers to switch category, instead of expanding sales in both categories. Following a similar logic, a price promotion of Citrus Hill refrigerated juice appears to be more beneficial than that of frozen juice, because it has a positive effect on own brand sales and in general negative effects on competitors' sales in the other category, which is not the case for a promotion of frozen juice.

The manufacturer of FloridaGold appears to be in a flexible position to maneuver regular price and promotion in both categories. A price promotion or regular price reduction of FloridaGold has beneficial effect on the brand's sales in both categories and generally hurts sales of competitors in both categories. Given its already low regular price level, this brand may want resort more to temporary price discounts. This seems to be corroborated by the actual practice of the company: FloridaGold's average price discount level is the highest in the refrigerated juice category and the second highest in the frozen juice category.

Minute Maid is a first-tier national brand in both categories in terms of price. It appears that there is little cross-category cannibalization resulting from a regular price reduction or price promotion of frozen juice. On the contrary, a price reduction or promotion for refrigerated juice attracts Minute Maid frozen juice consumers to switch category substantially, while having a much smaller effect on competitors' consumers category switching. Therefore, everything else being the same, Minute Maid would be better off by offering price reduction or promotion for the frozen juice rather than the refrigerated juice.

Tropicana has the largest market share and the highest price in the refrigerated orange juice category and the second largest market share in the frozen orange juice category. A price reduction or promotion of this brand has strong beneficial effects within each category, and there seems very little own brand cross-category impact. Therefore, the manufacturer of this brand should make her pricing and promotion decisions mainly based on the competitive situation within each category. It is worth pointing out though, that Tropicana's price reductions and promotions have beneficial effects on Dominick's sales in the other category, which is possibly caused by the brand's ability to draw store traffic. In addition, it appears that, among the four national brands, Tropicana's price reduction and promotion have the smallest negative effects on Dominick's sales within the same category and the most beneficial effects on Dominick's sales across categories. Therefore, the retailer should be more supportive to Tropicana than other companies operating in the orange juice category, for example by passing through Tropicana's regular price reductions. Tropicana may wish to utilize this information to strengthen its bargaining power with the retail chain.

DISCUSSION

Manufacturers cannot ignore the effects of their marketing activities in one category on the sales of their products in a related category. Jointly analyzing sales for individual brands in related categories provides substantive insights that may help manufacturers coordinate their marketing efforts across the categories in question. From a practical perspective, a methodology to estimate cross-category market response models *at the brand level*, as the one we have proposed in this study, is needed to support such decisions. Previous research on market-level effects of promotions has focused on either *within-category* analysis, or cross-category impact of price and promotions *at the category level*. The estimates obtained from our model indicate that brands exhibit quite distinctive patterns of cross-category effects and thus imply different cross-category pricing and promotion strategies/tactics for various manufacturers.

Results of our model provide a basis for discovering opportunities of positive cross-category impact and minimizing damage of negative cross-category impact through brand spill-over, category spill-over, store traffic, and budget constraint effects, in addition to more commonly assumed effects of product substitutability/complementarity. From a manufacturer's perspective, our approach enables one to examine whether it is more beneficial to reduce regular price only, offer price promotions while keeping regular price constant, or offer price promotion while increasing regular price for each category (Dhar and Raju 1998), as illustrated in our application. Further, as suggested by Walters (1991), manufacturers could utilize results generated by our approach to analyze the effect of trade promotions on a retailer's sales in related categories. Manufacturers that create positive cross-effects of their brands on other categories, for example by causing category spill over or generating store traffic, can use this knowledge to negotiate better trade promotion deals with retailers, and thus gain power in the channel relationship. Our analysis indicates that Tropicana is such an example.

The objective of our study is to assess all individual effects of a brand's marketing mix on sales of brands in another category. A limitation of our proposed methodology is that if it is applied to a larger number of categories, the number of parameters increases rapidly, potentially at the cost of predictive fit. Our primary objective, however, is not to develop a model that is parsimonious in estimating cross-elasticities for many categories, but rather, we intend to develop a model that, for a limited set of categories, estimates *all* within and between category cross-elasticities at the brand level, without imposing structural constraints on the parameters. Models that constrain these effects are more easily applicable to larger numbers of categories, but are more limited in their practical value to brand management. Although our model can in principle be applied to more than two categories, it would involve a very large number of parameters to be estimated and interpreted. Future research could be directed at developing more parsimonious model specifications that allow estimating own- and cross-effects with minimal amount of structural constraints, while at the same time enabling the analysis of a larger number of categories.

Another limitation of our model is that it is formulated at the brand level, rather than the SKU level. Brand management decisions in practice are sometimes made at the SKU level (Bucklin and Gupta 1999). Unfortunately, to our knowledge, no academic cross-category studies to date have provided an approach to dealing with the potentially very large number of SKU's in several categories. A plausible route for future research is to develop models along the lines of the methodology proposed by Fader and Hardie (1996) for single category data. Another important direction for future research is to explicitly decompose the overall marketing effects into the causal mechanisms that we have proposed in this study, and to measure and test each component while controlling for others. Song and Chintagunta (2002) have already provided an important step in that direction.

REFERENCES

Ainslie, Andrew and Peter E. Rossi (1998), "Similarities in Choice Behavior Across Product Categories," *Marketing Science*, 17 (2), 91-106.

Ansari, Asim, and Carl F. Mela (2002), "E-Customization," *Working Paper*, Columbia University.

Ansari, Asim and Rhaguram Iyengar (2002), "Semi-Parametric Bayesian Models for Binary and Polytomous Data," *Working Paper*, Columbia University.

Allenby, Greg M., Rossi, Peter E. (1999), "Marketing models of heterogeneity," *Journal of Econometrics*, 89, 57-78.

Andrews, Rick L., Asim Ansari and Imran S. Currim (2002), "Hierarchical Bayes versus Finite Mixture Conjoint Analysis Models: A Comparison of Fit, Prediction, and Partworth Recovery," *Journal of Marketing Research*, 39 (May), 87-98.

Bell, David R. and James M. Lattin (1998), "Shopping Behavior and Consumer Response to Retail Price Format: Why 'Large Basket' Shoppers Prefer EDLP," *Marketing Science*, 17(1), 66-88.

_______ and Jeongwen Chiang (2002), "Coupon Use and the Channel: A Within and Trans-Category Model of Shopper Behavior," *working paper*, University of Pennsylvania.

Blattberg, Robert C. and Kenneth Wisniewski (1988), "Issues in Modeling Store Level Scanner Data," *Working Paper*, No. 43, Center for Research in Marketing, The University of Chicago.

_______ and Edward George (1991), "Shrinkage Estimation of Price and Promotional Elasticities: Seemingly Unrelated Equations," *Journal of the American Statistical Association*, 86, 304-315.

Boatwright, Peter, Robert McCulloch and Peter Rossi (1999), "Account Level Modeling for a Trade promotion: An Application of a Constrained parameter Hierarchical Model," *Journal of the American Statistical Association*, 94 (448), 1063-1073.

Bucklin, Randolph E. and Sunil Gupta (1999), "Commercial Use of UPC Scanner Data: Industry and Academic Perspectives," *Marketing Science*, 18 (3), 247-273.

Chintagunta, Pradeep and Sudeep Haldar (1998), "Investigating Purchase Timing Behavior in Two Related product Categories," *Journal of Marketing Research*, 35 (February), 43-53.

Dhar, Sanjay K. and Jagmohan S. Raju (1998), "The Effects of Cross-Ruff Coupons on Sales and Profits," *Management Science*, 44 (11), 1501-1516.

Erdem, Tulin (1998), "An Empirical Analysis of Umbrella Branding," *Journal of Marketing Research*, 35 (August), 339-351.

_______ and Russell S. Winer (1999), "Econometric Modeling of Competition: A Multi-Category Choice-Based Mapping," *Journal of Econometrics*, 89(1/2), 159-175.

Escobar, Michael D. (1994), "Estimating Normal Means with A Dirichlet Process Prior," *Journal of the American Statistical Association*, 89 (425), 268-277.

_______ and Mike West (1995), "Bayesian Density Estimation and Inference Using Mixtures," *Journal of the American Statistical Association*, 90 (430), 577-588.

Fader, Peter S. and Bruce G.S. Hardie (1996), "Modeling Consumer Choice Among SKUs," *Journal of Marketing Research*, 33 (November), 442-452.

Fan, Jinqing and Irene Gijbels (1996). *Local Polyonmial Regression and Its Applications*. London: Chapman and Hall.

Ferguson, T.S. (1973), "A Bayesian Analysis of Some Nonparametric Problems," *The Annals of Statistics*, 1, 209-230.

Foekens, Eijte W., PeterS.H. Leeflang, and Dick R. Wittink (1994), "A Comparison and an Exploration of the Forecasting Accuracy of a Nonlinear Model at Different Levels of Aggregation," *International Journal of Forecasting*, 10 (2), 245-61.

Gupta, Sachin, Pradeep Chintagunta, Anil Kaul, and Dick R. Wittink (1996), "Do Household

Scanner Data Provide Representative Inferences from Brand Choice: A Comparison with Store Data," *Journal of Marketing Research*, 33 (4), 383-398.

Heerde, Harald Van, Peter S.H. Leeflang and Dick R. Wittink (2001), "Semiparametric Analysis to Estimate the Deal-Effect Curve," *Journal of Marketing Research*, 38 (May), 197-215.

Hoch, Stephen J., Byung-Do Kim, Alan L. Montgomery, and Peter E. Rossi (1995), "Determinants of Store-level Price Elasticities," *Journal of Marketing Research*, 32 (1), 17-29.

Kalyanam, Kirthi and Thomas S. Shively (1998), "Estimating Irregular Pricing Effects: A Stochastic Spline Approach," *Journal of Marketing Research*, 35 (February), 16-29.

Kamakura, Wagner A. and Gary J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26 (November), 379-390.

Kopalle, Praveen K., Carl F. Mela, and Lawrence Marsh (1999), "The Dynamic Effects of Discounting on Sales: Empirical Analysis and Normative Pricing Implications," *Marketing Science*, 18 (3), 317-332.

Leeflang, Peter S.H. and Dick R. Wittink (2000), "Building Models for Marketing Decisions: Past, Present and Future," *International Journal of Research in Marketing*, 17 (2/3), 105-126.

Manchanda, Puneet, Asim Ansari, and Sunil Gupta (1999), "The 'Shopping Basket': A Model for Multicategory Purchase Incidence Decisions," *Marketing Science*, 18(2), 95-114.

Seetharaman, P. B., Andrew Ainslie, and Pradeep K. Chintagunta (1999), "Investigating Household State Dependence Effects Across Categories," *Journal of Marketing Research,* 36 (November), 488-500.

Russell, Gary J. and Wagner A Kamakura (1997), "Modeling Multiple Category Brand Preference with Household Basket Data," *Journal of Retailing*, 73(4), 439-461.

_______, S. Ratneshwar, Allan D. Shocker *et al.* (1999), "Multiple-Category Decision-Making: Review and Synthesis," *Marketing Letters*, Special Issue on the Fourth International Choice Symposium, 10 (3), 317-330.

_______ and Ann Petersen (2000), "Analysis of Cross-category Dependence in Market Basket Selection," *Journal of Retailing*, 76 (3), 367-392.

Song, Inseong and Pradeep Chintagunta (2001), "Investigating Cross-category Effects of Retailer's Marketing Activities: Application of a Random Coefficients Choice Model with Aggregate Data," *Working Paper University of Chicago.*

Walters, Rockney G. (1988), "Retail Promotions and Retail Store Performance: A Test of Some Key Hypotheses," *Journal of Retailing*, 64 (2), 153-180.

_______ (1991), "Assessing the Impact of Retail Price Promotions on Product Substitution, Complementary Purchase, and Interstore Sales Displacement," *Journal of Marketing*, 55 (April), 17-28.

Wittink, Dick R., M. J. Addona, W. J. Hawkes, and J. C. Porter (1988), "SCAN*PRO: The Estimation, Validation and Use of Promotional Effects Based on Scanner Data," *Working Paper,* Cornell University.

Zellner, Arnold (1972), *An Introduction to Bayesian Inference in Econometrics*. New York: Wiley.

Table 1

CROSS-CATEGORY EFFECTS OF PRICE PROMOTION ON SALES

Brand in Another Category:	Own Brand	Other Brands
Substitutability/complementarity	-/+	-/+
Brand spill-over	+	-
Category spill-over	+	+
Store traffic	+	+
Budget constraint	+	+

Table 2
DESCRIPTIVE STATISTICS OF THE ORANGE JUICE DATA

Frozen Orange Juice

Brands	Weekly Sales (ounces)		Regular Price (cents/oz)		Price Discount (cents/oz)	
	Mean	SE	Mean	SE	Mean	SE
Citrus Hill	1476.4	3105.1	12.21	1.23	0.38	1.05
FloridaGold	1725.0	3494.6	11.96	1.12	0.44	1.03
MinuteMaid	3187.6	3859.8	13.34	1.04	0.34	0.84
Tropicana	3410.6	5467.7	12.21	0.73	0.56	1.07
Dominick's	5256.8	6416.1	10.93	0.74	0.10	0.30

Refrigerated Orange Juice

Brands	Weekly Sales (ounces)		Regular Price (cents/oz)		Price Discount (cents/oz)	
	Mean	SE	Mean	SE	Mean	SE
Citrus Hill	9622.5	20290.0	3.64	0.32	0.07	0.16
FloridaGold	2194.3	3571.0	3.38	0.45	0.19	0.40
MinuteMaid	26519.6	38104.0	4.60	0.25	0.09	0.15
Tropicana	39646.1	28384.7	4.72	0.35	0.16	0.18
Dominick's	30431.3	40360.2	2.78	0.40	0.12	0.23

Table 3
EFFECTS OF STORE DEMOGRAPHICS[1]

Frozen Juice	*Citrus Hill*	*FloridaGold*	*Minute Maid*	*Tropicana*	*Dominick's*
Intercept	**7.426**	**7.706**	**7.265**	**6.943**	**13.018**
age	0.002	0.004	-0.019	0.000	-0.006
education	**0.341**	**0.319**	**0.324**	**0.253**	**0.248**
income	-0.126	-0.127	-0.125	-0.118	-0.131
hhsize	0.086	0.095	0.025	0.058	0.047
sinhouse	0.002	-0.012	0.067	-0.001	0.045
density	-0.064	-0.072	-0.069	-0.075	-0.086
Refrigerated Juice	*Citrus Hill*	*FloridaGold*	*Minute Maid*	*Tropicana*	*Dominick's*
Intercept	**10.229**	**4.360**	**16.917**	**8.358**	**20.400**
age	0.091	**0.136**	0.085	**0.202**	0.099
education	0.115	0.122	0.091	**0.216**	**0.163**
income	-0.048	-0.108	-0.028	-0.041	**-0.239**
hhsize	**0.119**	**0.118**	0.088	**0.142**	**0.185**
sinhouse	-0.080	-0.082	-0.085	-0.166	-0.065
density	-0.062	-0.073	-0.065	**-0.114**	**-0.123**

[1] Boldface font indicates that the posterior 95% credible interval does not cover zero

Table 4
POSTERIOR MEANS OF POOLED WITHIN-
AND BETWEEN-CATEGORY EFFECTS AND THEIR PRECISIONS[1]

	Within Category			*Between Category*	
Frozen Juice	*Own*	*Cross*	*Precision*	*Cross Effect of Refrigerated*	*Precision*
Regular Price	**-3.37**	**1.23**	**0.68**	-0.23	**1.01**
Price Discount	**1.62**	**-0.34**	**16.52**	-0.04	**14.36**
	Within Category			*Between Category*	
Refrigerated Juice	*Own*	*Cross*	*Precision*	*Cross Effect of Frozen*	*Precision*
Regular Price	**-2.96**	**0.84**	**0.36**	0.01	**1.32**
Price Discount	**2.60**	**-0.44**	**4.60**	0.00	**10.97**

[1] Boldface font indicates that the posterior 95% credible interval does not cover zero

Table 5

POSTERIOR MEAN INDIVIDUAL WITHIN-CATEGORY EFFECTS[1,2]

Frozen Juice Sales: Regular Price

	Citrus Hill	*FloridaGold*	*MinuteMaid*	*Tropicana*	*Dominick's*
Citrus Hill	**-4.77**	**0.76**	**2.45**	0.55	**0.93**
FloridaGold	0.25	**-6.29**	**1.43**	**2.37**	**1.70**
MinuteMaid	**0.94**	**0.79**	**-3.98**	**1.55**	**2.61**
Tropicana	**0.98**	**0.37**	**1.44**	**-4.36**	**2.34**
Dominick's	**1.16**	**1.46**	**3.09**	**1.08**	**-7.90**

Refrigerated Juice Sales: Regular Price

	Citrus Hill	*FloridaGold*	*MinuteMaid*	*Tropicana*	*Dominick's*
Citrus Hill	**-2.30**	-0.03	-0.29	-0.27	**0.46**
FloridaGold	**1.61**	**-3.23**	0.42	-0.62	-0.25
MinuteMaid	**1.07**	**0.26**	**-10.09**	**2.86**	**1.30**
Tropicana	**-0.79**	0.16	**3.51**	**-4.07**	**-0.45**
Dominick's	**1.02**	-0.10	**1.53**	-0.52	**-2.43**

Frozen Juice Sales: Price Discount

	Citrus Hill	*FloridaGold*	*MinuteMaid*	*Tropicana*	*Dominick's*
Citrus Hill	**1.93**	**-0.23**	**-0.40**	**-0.33**	**-0.42**
FloridaGold	**-0.07**	**1.86**	**-0.18**	**-0.36**	-0.12
MinuteMaid	**-0.26**	**-0.21**	**1.19**	**-0.25**	**-0.47**
Tropicana	**-0.34**	**-0.24**	**-0.48**	**1.69**	**-0.52**
Dominick's	**-0.26**	**-0.51**	**-0.49**	**-0.63**	**1.76**

Refrigerated Juice Sales: Price Discount

	Citrus Hill	*FloridaGold*	*MinuteMaid*	*Tropicana*	*Dominick's*
Citrus Hill	**4.16**	**-0.21**	-0.17	**-0.96**	**-0.19**
FloridaGold	**-0.38**	**2.18**	**-0.48**	0.21	**-0.46**
MinuteMaid	**-0.55**	**-0.11**	**3.37**	**-0.65**	**-0.73**
Tropicana	**-0.81**	**-0.11**	**-0.24**	**2.14**	-0.15
Dominick's	**-0.37**	**-0.30**	**-0.71**	**-0.30**	**2.98**

[1] Boldface font indicates that the posterior 95% credible interval does not cover zero

[2] Cell entries indicate the effect of the price of the brand in the column on the sales of the brand in the row.

Table 6

POSTERIOR MEAN INDIVIDUAL BETWEEN CATEGORY EFFECTS[1]

Frozen Juice Sales: Regular Price Effect Refrigerated Juice

	Citrus Hill	*FloridaGold*	*MinuteMaid*	*Tropicana*	*Dominick's*
Citrus Hill	**1.35**	**-0.27**	0.64	**-2.55**	**0.46**
FloridaGold	**-0.83**	**-0.30**	0.21	0.66	-0.29
MinuteMaid	-0.09	-0.01	**-2.91**	0.25	**0.58**
Tropicana	-0.49	0.03	**-1.23**	0.40	**0.35**
Dominick's	0.50	**0.38**	-0.35	**-2.26**	0.09

Refrigerated Juice Sales: Regular Price Effect Frozen Juice

	Citrus Hill	*FloridaGold*	*MinuteMaid*	*Tropicana*	*Dominick's*
Citrus Hill	**-0.54**	0.03	**1.37**	**0.69**	**-0.85**
FloridaGold	**1.57**	**-0.83**	0.52	0.07	0.59
MinuteMaid	-0.25	**-0.49**	0.07	**0.53**	**0.53**
Tropicana	-0.09	**0.61**	**0.92**	**0.89**	**-0.70**
Dominick's	-0.04	**0.35**	**-1.54**	**-2.47**	**-0.64**

Frozen Juice Sales: Price Discount Effect Refrigerated Juice

	Citrus Hill	*FloridaGold*	*MinuteMaid*	*Tropicana*	*Dominick's*
Citrus Hill	0.18	0.06	0.02	0.02	**-0.33**
FloridaGold	**-0.35**	**0.20**	**-0.27**	0.01	0.13
MinuteMaid	-0.10	**-0.11**	**-0.37**	-0.15	0.06
Tropicana	0.02	-0.03	-0.12	-0.09	**0.34**
Dominick's	-0.18	0.00	0.10	0.18	-0.15

Refrigerated Juice Sales: Price Discount Effect Frozen Juice

	Citrus Hill	*FloridaGold*	*MinuteMaid*	*Tropicana*	*Dominick's*
Citrus Hill	0.03	-0.05	**-0.18**	0.00	0.01
FloridaGold	**0.20**	**0.41**	**0.22**	**-0.08**	-0.17
MinuteMaid	**0.27**	**-0.19**	-0.03	**-0.11**	**-0.72**
Tropicana	-0.05	**-0.27**	**-0.20**	0.04	**0.31**
Dominick's	**0.11**	**0.56**	-0.01	**0.12**	**-0.29**

[1] Boldface font indicates that the posterior 95% credible interval does not cover zero.

[2] Cell entries indicate the effect of the price of the brand in the column on the sales of the brand in the row

Table 7

RESIDUAL LOG-SALES CORRELATIONS[1,2]

		Frozen *Citrus*	*Florida*	*Minute*	*Tropic*	*Domin*	*Refrig* *Citrus*	*Florida*	*Minute*	*Tropic*	*Domin*
Frozen	*Citrus Hill*	**0.58**									
	FloridaGold	**0.11**	**0.63**								
	MinuteMaid	0.11	0.07	**0.44**							
	Tropicana	**0.17**	0.10	0.11	**0.47**						
	Dominick's	0.00	0.10	0.07	0.04	**0.47**					
Refrig	*Citrus Hill*	**0.18**	0.06	0.04	**0.23**	0.10	**0.42**				
	FloridaGold	0.01	**0.08**	**0.08**	0.03	-0.01	0.01	**0.89**			
	MinuteMaid	0.04	0.05	**0.18**	0.09	0.07	0.06	-0.04	**0.45**		
	Tropicana	0.07	0.03	0.03	**0.16**	**0.19**	0.09	**-0.11**	0.07	**0.42**	
	Dominick's	**0.15**	0.06	**0.14**	0.08	-0.06	**0.15**	-0.04	-0.05	-0.04	**0.50**

[1] Boldface font indicates that the posterior 95% credible interval does not cover zero.

[2] Residual SE's on the diagonal.

Figure 1
SMOOTH CATEGORY TIME-SALES PLOTS

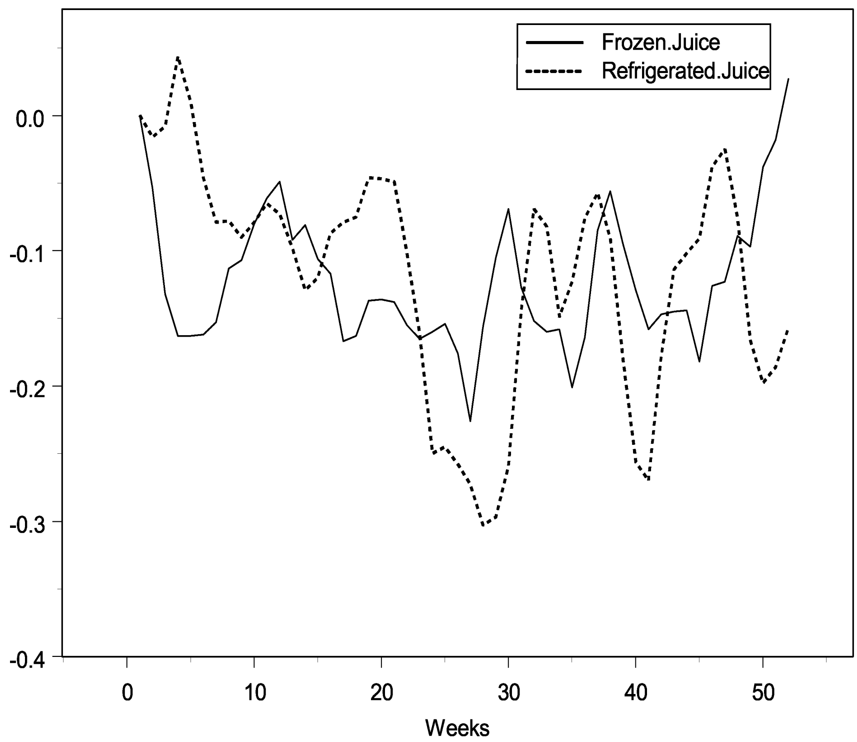
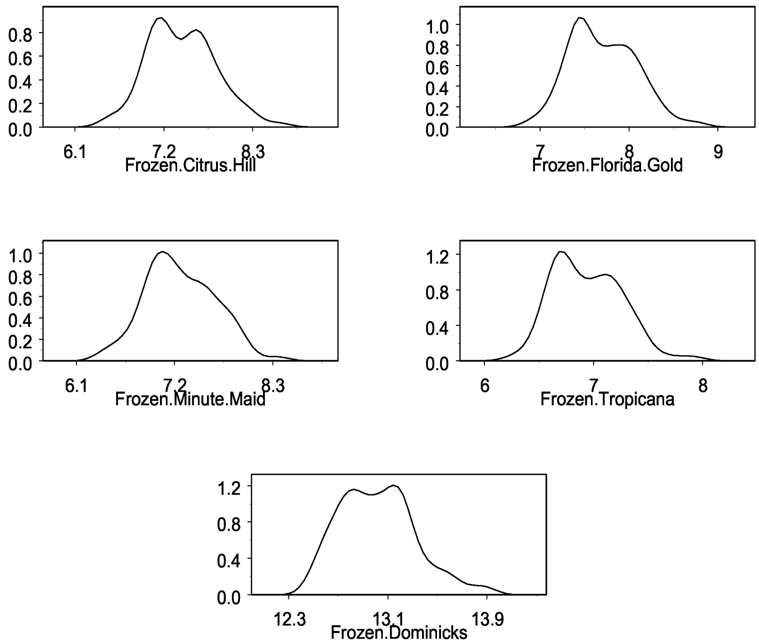


Figure 2

DPP BASED DENSITY PLOTS FOR BRAND CONSTANTS

Frozen Juice



Refrigerated Juice