A Model-Based Approach to Setting Optimal Retail Markups

Michel Wedel, Jie Zhang and Fred Feinberg

February, 2004

University of Michigan Business School, 701 Tappan Street, Ann Arbor, Michigan 48109, USA.

A Model-Based Approach to Setting Optimal Retail Markups

SUMMARY

This study provides a model-based approach for implementing retail category management with a focus on mark-up decisions. We develop a reduced form Bayesian econometric model that captures price and sales variation for a retail chain. Based on this model, we derive the optimal retail markups on individual brands within a product category by maximizing the total category profit over a finite time horizon in a Bayesian framework. We demonstrate our approach using store-level weekly sales data from a retail chain. Our empirical findings indicate that, contrary to common practice in the retail industry, markups on high share brands could well be greater. We also find that category profit can be increased by driving down wholesale prices via competitive force, without an increase in overall category demand. We discuss the implications of our approach for retailers and other channel partners.
.

KEYWORDS: Bayesian Pooling, Nonparametric Smoothing, Global Optimization, Category Management.

1 INTRODUCTION

For retailers, category management (CM) is the process of maximizing *category* profits through coordinated item assortment, promotion and pricing. It has been gaining popularity because retailers have come to realize that a better understanding of the interdependencies in consumer and manufacturer behavior among various brands in a category can increase the profitability of the category as a whole.[1] CM is a significant departure from the traditional approach that considers brands independently, and has been adopted by an increasing number of companies in the retail industry: 78% of department stores, 74% of discount stores, and 45% of supermarkets deploy category-management in some form according to a recent survey.[2] Findings by Zenor (1994) and Basuroy, Mantrala and Walters (2001) support the notion that adopting a CM perspective could increase total profit. Yet, despite the optimistic view, there is little consensus on how to implement CM in order to achieve its greatest profit potential. This is at least in part caused by a lack of analytical tools to support CM decisions, and calls for econometric methods that enable retail managers to implement CM.

The main objective of this study is therefore to offer a model-based approach for implementing CM in response to needs of the retail industry. We focus on a key aspect of retail CM in this study: pricing decisions. For a retailer, pricing decisions mainly involve setting markups over the wholesale prices of the brands in the categories it sells. We develop a reduced form econometric model that captures store-level price and sales variation in a

[1] *Progressive Grocer*, December 1993

[2] *Chain Store Age*, March 2000

comprehensive manner, and then provide a global optimization approach to deriving the optimal markups for a category.

Neoclassical double-log demand systems, akin to Tobin's (1950), have been a popular choice for modeling demand from cross sectional and time series data (cf. De Crombrugghe, Palm and Urbain 1997). For example, extending earlier work of Blattberg and George (1991), Montgomery (1997) develops a Hierarchical Bayes model for the effects of marketing variables on log-sales. Kopalle, Mela and Marsh (1999) extend the standard double-log neoclassical demand function to include dynamic effects, using a varying parameters specification. Building on those approaches, we intend to offer the following contributions in this study. In the demand functions, we allow price effects to vary among individual brands, and accommodate sales heterogeneity across stores, time-varying trends in sales levels, and price endogeneity. Based on parameter estimates from this model, we adopt a global optimization procedure in the Bayesian framework and maximize the category profit over a given time horizon. A challenge here is that standard log-log demand functions do not generally yield interior solutions to maximizing a retailer's category profit. We propose to include a price-index in the demand model, similar in spirit to that used in the AIDS model (Deaton and Muellbauer 1980), which allows for interior solutions to the optimal category pricing problem. Our optimization method explores the posterior category profit surface, which resolves the problem of obtaining corner solutions for optimal prices from the standard log-log or semi-log demand functions with constraints. Using this approach, we obtain optimal retail markups on all brands in the category.

We proceed by detailing our model, describing the optimization procedure, illustrating the application of our approach using data for the toothbrush category, and discussing the managerial implications of our findings.

2 MODEL AND ESTIMATION

2.1 Model Description

We develop a hierarchical Bayesian model for store level weekly sales data. Let $r = 1, ..., R$ indicate stores, $i = 1, ...N$ indicate brands, and $t = 1, ..., T$ indicate weeks. The model entails a simultaneous equation specification for demand and prices (cf. Kloek and Van Dijk 1978). The first equation describes store-level sales as a function of the prices of the brands in the category, as well as a category price index and time trends. The second describes prices as a function of *observed* last-period retail prices and sales volumes. Both are reduced form, Cobb-Douglas type equations. Letting $q_{r,t} = [q_{r,t,i}]$ and $p_{r,t} = [p_{r,t,i}]$ be vectors of sales quantities and retail prices for store r in week t, we have:

$$\ln(q_{r,t}) = \mu_r + \Gamma \ln(p_{r,t}) + \beta\, s_{r,t} + \lambda(t) + \epsilon_{r,t} \quad (1)$$

$$\ln(p_{r,t}) = \nu_r + A \ln(p_{r,t-1}) + \Psi \ln(q_{r,t-1}) + \delta_{r,t}, \quad (2)$$

where $\mu_r = [\mu_{r,i}]$ and $\nu_r = [\nu_{r,i}]$ are $N \times 1$ vectors of store-specific constants in the sales and pricing functions, respectively; $\Gamma = [\gamma_{i,j}]$, $A = [\alpha_{i,j}]$, and $\Psi = diag[\psi_{i,i}]$ are $N \times N$ matrices of price coefficients on demand, and price-reaction and lagged sales coefficients on pricing, respectively. $s_{r,t}$ is the Stone price index (Stone 1954), computed as the weighted average price in store r and week t,[3] and $\beta = [\beta_i]$ is a vector of the coefficients of the price index on each brand's sales. The expected sign for each of the β_i's is negative, allowing for an

[3] The weights are computed from an initialization period and thus do not vary with time.

interior solution for optimal retail prices, as will be explained in more detail shortly.

Our reduced form pricing equations include own and competitive prices, the effects of which are reflected in the coefficients Γ, as well as own last-period sales, as captured by Ψ (for a similar formulation, see, for example, Yang, Chen and Allenby 2003). We specify store-specific constants, μ_r and ν_r, to capture differences in baseline brand sales and prices across stores. We represent store heterogeneity through a random effects specification:

$$\mu_r \sim N\left(m, V_\mu\right), \quad \nu_r \sim N\left(n, V_\nu\right), \tag{3}$$

where $m = [m_i]$, $n = [n_i]$ are $(N \times 1)$ vectors and V_μ and V_ν are $(N \times N)$ matrices. Multivariate Normal $N\left(0, 10^5 \times I\right)$ priors are specified for m, and n, and inverse-Wishart priors $IW\left(N+3, I\right)$ are used for V_μ and V_ν.

Direct estimation of the own- and cross-effects (Γ) in demand equations often gives rise to instability of the estimates due to the large number of cross-price effects to be estimated ($N^2 - N$ coefficients for N brands). To avoid this problem, we employ a Bayesian shrinkage method, considering the individual own- and cross-price effects in each equation random draws from their category level means. The specification is as follows:

$$vec\left(\Gamma\right) \sim N_0^\infty\left(vec\left(g_o I + g_c\left(\iota\iota' - I\right)\right), V_\gamma \times I\right), \tag{4}$$

with I an identity matrix of appropriate dimensionality, ι a column vector of ones and g_o, g_c and V_γ scalars. Hierarchical Bayesian estimates of the individual price effects (Γ) borrow stability from the category-average own (g_o) and cross-effect coefficients (g_c). The variance of these effects is denoted by V_γ. In addition, the cross-effects are constrained to be positive by drawing them from the normal distribution in (4) which is MVN truncated below at zero (see Boatwright, McCulloch and Rossi 1999). Note that the single-point priors themselves would correspond to a model where only aggregate own- and cross-effects are estimated.

We pool the price function coefficients in a similar manner, but do not constrain the individual coefficients (A) since we do not have solid theory on their directions:

$$vec\left(A\right) \sim N\left(vec\left(a_o I + a_c\left(\iota\iota' - I\right)\right), V_\alpha \times I\right), \tag{5}$$

where a_o and a_c are category-average own- and cross-coefficients, and V_α is the variance term

Mildly informative priors are specified for the means: $(g_0, g_c)' \sim N\left((-2,1)', 0.1 \times I\right), (a_0, a_c)' \sim N\left((0.5, 0.5)', 0.1 \times I\right)$.[4] We use prior scalar variances: $V_\gamma^{-1} \sim W\left(2,1\right)$and $V_\alpha^{-1} \sim W\left(2,1\right)$. We use standard uninformative priors for $\Psi \sim N\left(0, 10^5 \times I\right)$ and a truncated Normal prior for the price index coefficients $\beta \sim N_{-\infty}^{0}\left(0, 10^5 \times I\right)$, which ensures that no combination of prices rising to infinity can possibly drive total category demand to infinity, and thus enabling interior solutions to the category pricing problem.

We capture time-variation in brand sales through a smooth vector function: $\lambda\left(t\right) = \left[\lambda_i\left(t\right)\right]$, using a Hierarchical Bayes extension of local polynomial regression (Fan and Gijbels 1997). This approach is based on the assumption that the time-sales function can be locally approximated in a neighborhood $(-h, h)$ of time point t by a Taylor series expansion:

$$\lambda_i\left(t\right) \approx \sum_{p=0}^{P-1} \int_{t-h}^{t+h} \left(z-t\right)^p \Omega_i\left(dz\right). \tag{6}$$

Equation (6) describes $\lambda\left(t\right)$ locally by $\left(P-1\right)^{th}$-order local polynomials in the neighborhood h. The coefficients of the local polynomial change stochastically over time according to a P-

[4] The priors are derived from previous studies. Analyses reveal that the posterior distributions are relatively insensitive to the specific prior chosen.

variate Wiener process with drift: $\Omega_i(s) = \omega_i + \Delta' B_i(s)$, where $B_i(s)$ is standard P-variate Brownian motion, ω_i is a $(P \times 1)$ vector of means, and $V_\omega = \Delta'\Delta$ a $(P \times P)$ covariance matrix. Local polynomial regression offers good min-max efficiency properties and has negligible boundary effects. We use polynomials of order $P = 2$, a global bandwidth of $h = 13$, and a uniform kernel in our empirical application (cf. Fan and Gijbels 1997). The following priors are employed: $IW(h+3, I)$ for V_ω and $N\left(0, 10^5 \times I\right)$ for ω_i.

Finally, we account for price endogeneity due to unobserved causes by assuming a multivariate normal distribution for the error terms in the sales and price functions, $\vartheta_{r,t} = (\epsilon_{r,t,i}, \delta_{r,t,i})$, stacked. Our procedure allows the covariance of the sales and pricing errors in the full model (Equations 1 and 2) to be formulated according to a SUR specification: $\vartheta_{r,t} \sim N(0, \Sigma)$. A conjugate $IW(2N+3, I)$ prior is used for Σ. The model can thus be seen as an extended hierarchical version of the Bayesian multivariate regression model (Zellner 1972, p. 224-246).

2.2 Estimation

We employ Markov Chain Monte Carlo (MCMC) methods to estimate the model. All prior distributions are standard conjugate distributions. In each iteration, s, of the MCMC chain, we draw from the full conditional distribution of a block of parameters, conditional on the values of the other parameters obtained from the last draw: $\Xi_s \mid \Xi_{/s}, D$, where $D = \{\ln(q_{r,t}), \ln(p_{r,t})\}$ denotes the data and Ξ is the collection of all model parameters. We use 10,000 draws and a burn-in of 5,000. The algorithm is started from approximate GLS estimates and convergence is monitored through plots of key parameters against iterates.

Results of synthetic data analyses indicated that the algorithm performs well in recovering all true parameters and converges to a stationary distribution well before 5000 draws.

3 DERIVING THE OPTIMAL RETAIL MARKUPS

The category management approach requires a retailer to coordinate decisions on all brands in a category to maximize the total category profit. We choose to focus on how to set markups as it is one of the most important decisions for retailers. We adopt a Bayesian approach to decision theory to find the optimal settings for the control variables (Dorfman 1995, p. 88-96). Since retailers usually make pricing and promotion calendar decisions for a certain time period in advance, we compute the total category profit in T weeks to derive a constant unit markup for each brand over the planning horizon and across stores. Such constant markups are consistent with retailers' general practice and are easy to implement.

The objective function is category profit over a period of T weeks as given by:

$$\Pi\left(M;\Xi\right)=\sum_{t=1}^{T}\sum_{r=1}^{R}M'q_{r,t}\left(\Xi\right), \tag{7}$$

where $M=[M_i]$ is an $N\times 1$ vector of the retailer's unit markup of brand i, Ξ represents parameters in the model, and $q_{r,t}\left(\Xi\right)$ is the $N\times 1$ vector of sales quantities. In the empirical analysis, $q_{r,t}\left(\Xi\right)$ and $\Pi(M;\,\Xi)$ are computed as follows. Let $W_t=[W_{i,t}]$ be a vector of brand i's wholesale price at t, $C_r^*=[C_{r,i}^*]$ be a vector of the retailer's idiosyncratic unit variable cost of brand i at store r. Then, $p_{r,t}=W_t+M+C_r^*$ is the vector of retail prices observed in the data. Let $C_{r,t}=W_t+C_r^*$ be the retailer's total unit variable cost of brands at store r in week t. $C_{r,1}$, i.e., $C_{r,t}$ at the first week of the period of interest, are assumed known to the retailer. We obtain the expressions for price and sales at $t=1$ by substituting

$p_{r,1} = C_{r,1} + M$ into equation (1), initializing with $q_{r,0}$ as the observed sales in the last time period, and taking exponents. Then, using equations (1) and (2) recursively, we obtain $q_{r,t}(\Xi)$ for $t = 2, \ldots, T$ as a function of the parameters Ξ. T is determined by the retailer, and in practice is often set to three months (*i.e.*, 13 weeks). Therefore, we chose $T = 13$ weeks, which was also used by Neslin, Powell and Stone (1995).

We wish to maximize (8) to derive the optimal markups M^*. There are several complications in doing so through conventional methods. First of all, the standard log-log demand functions usually do not yield interior solutions for optimal category pricing. This arises in the presence of cross-effects (positive off-diagonal elements of Γ, which we observe in our empirical application): if the price of one brand goes to infinity, demand for that brand becomes zero, but the demand for one or more other brands is driven to infinity, which drives category demand and profit to infinity. Some authors have imposed constraints to alleviate this problem, but corner solutions may still arise if the cross-effects are non-negligible. In our model, the inclusion of the category price index $s_{r,t}$ in equation (1) eliminates this problem: because β is negative for each brand, as the price of a brand goes up, the linear price index effect eventually dominates the log brand price effects and drives category demand downward. It is thus impossible to set some brands at artificially high-prices to force category demand ever higher.

We adopt a global optimization method to maximize the category profit by exploring the posterior profit function through simulation. We draw a set of markups $M^{v,s} \sim U(0,1)$ for $v = 1, ..., V$, where s indicates each draw. In the empirical application, we choose $V = 1024$ which represents eight replications of a a 2^7 orthogonal array (Niederreiter 1992)), where 7 is the number of brands in the category. We normalize the markups so that the sum of them is equal to their sum in the previous three months (i.e., 13 weeks) in the data. While this constraint is not needed to obtain interior solutions to the category pricing problem, it is intuitively appealing in that it yields markups in the order of magnitude of previous values, and enables a fairer comparison with current practice and other heuristic price-setting methods. Quasi-Monte Carlo draws are used, which greatly improves the efficiency of the optimization. At each fifth target draw of the Gibbs sampler, s, we compute the 13-week category profit, $\Pi\left(M^{v,s}, \Xi^{s}\right)$ using (1) and (8) at the current parameter values, Ξ^{s}, and a draw from $\vartheta_{r,t}^{s} \sim N\left(0, \Sigma^{s}\right)$. We retain $M^{*,s} = \text{argmax}\left[\tilde{\Pi}\left(M^{v,s}\right); v = 1, ..., V\right]$, and obtain the optimal profit using (8) across the draws of the Gibbs sampler as: $\tilde{\Pi}\left(M^{*}\right) = \sum_{s=1}^{S} \Pi\left(M^{*,s}, \Xi^{s}\right)/S$. In addition to the optimal markups and profit, we also compute predicted sales, revenue, total cost and unit cost.

4 APPLICATION

4.1 Data Description

We use store-level scanner data on toothbrushes from the Dominick's chain in the Chicago market.[5] The data include weekly sales, shelf prices, and markups for each brand of the category in 76 stores and 66 weeks. We use one week for initializing the lagged price and sales variables, 52 weeks for estimation, and 13 weeks for holdout validation and optimization. There are seven brands: Aquafresh, Butler, Colgate, Crest, Oral B, Pepsodent, and Reach. Table I shows the means and standard deviations of sales, prices and markups (note that

[5]We are grateful to The Kilts Center for Marketing at the University of Chicago Graduate School of Business for making the data available for academic usage.

the markups show very little variation).

[INSERT TABLE I ABOUT HERE]

4.2 Model Estimation Results

We present the most relevant parameter estimates in several tables and figures. Figure 1 displays the brand-specific sales-time smooth functions ($\lambda(t)$), which reveals a differentially evolving sales pattern for the brands in the category, and shows seasonal variation and possibly holiday-related fluctuations. Note that the trends captured here are net of own- and cross-price effects, changes in the overall category price level, and differences among stores. Table II shows the posterior mean of the brand specific sales and price constants (m and n) and their standard deviations ($diag\left(V_{\mu}^{1/2}\right)$ and $diag\left(V_{\nu}^{1/2}\right)$).

[INSERT FIGURE 1 AND TABLE II ABOUT HERE]

Table III shows the estimates of the pooled sales and pricing equation parameters (g_0, g_c, a_0 and a_c). These can be seen as category averages of the individual coefficients in the demand and price equations. The posterior mean of the own-price autoregressive coefficient is 0.47 and is significant,[6] which suggests that there is significant autocorrelation of prices at the category level. As to the price effects on sales, the posterior mean own-price demand effect is -4.35, and the cross-price effect is 0.49, both being significant. Thus, the brands in this category appear to be quite price-elastic and there is considerable competition among them.

[6] For ease of exposition we will call an effect "significant" if its posterior credible interval does not cover zero.

Table IV presents the individual own- and cross-price coefficients (Γ) and the category price index effects (β) in the demand functions. Two brands, Butler and Pepsodent, show significant price index effects. The magnitude of these two effects is large, indicating that the brands' sales go down substantially as the average price level of the category increases. Interestingly, Butler and Pepsodent are the two lowest-priced brands in the category, suggesting that low-priced items may be more prone to sales fluctuations caused by changes in the category price level. Looking at the matrix Γ, a few own-price coefficients are fairly large, such as for Aquafresh, Oral B and Reach, which reveals that demand for these brands is fairly elastic. The cross-price coefficients suggest asymmetric competition among brands. Higher-priced brands, such as Colgate, Oral B, and Reach seem to have more clout over others, while low-priced brands such as Butler and Pepsodent appear to be more vulnerable to other brands' influences.

[**INSERT TABLES III AND IV ABOUT HERE**]

Table V presents the posterior mean of the coefficients in the price equations (A and Ψ). We first focus on Ψ, which represents the effect of own-lagged-sales on price. This effect is positive and significant for six out of the seven brands (with the exception of Aquafresh, for which the coefficient is positive but insignificant), and the coefficients are quite large in magnitude. Parameters in A represent lagged own- and cross-price effects. All of the own-lagged-price effects are positive and significant, as expected. In general, these coefficients are larger than those of lagged-prices of the other brands. The latter indicate asymmetric competitive price reactions in this category. The most common pattern seems to be a leader-follower type, indicated by a strong positive lagged-price coefficient of one brand on the other, but a minimal effect the other way around. It appears that the leader in these pairs is mostly a higher-priced brand and the follower a lower-priced brand.

[INSERT TABLE V ABOUT HERE]

We use thirteen-week price and sales data for hold-out validation of the model, which corresponds to the period over which we derive the optimal markups. The prices are predicted with a holdout correlation of 0.60 (RMSE = 0.024), and the sales equation yields a holdout correlation of 0.87 (RMSE = 0.527). There are some differences among brands, especially with respect to the price forecasts, but the overall predictive performance of the model in the holdout period can be considered satisfactory.

4.3 Optimization Results

For purposes of comparison, we evaluate our optimal markups versus those derived from three heuristic alternatives which may be used by retailers. We apply the same constraint to the markup allocation for these heuristics and fix the total markup over all brands to the actual level in the last thirteen weeks of the estimation data. This is a reasonable constraint from a retailer's perspective, and it enables a more clear-cut comparison since the total markup for different procedures is fixed to be the same. The three heuristic procedures considered are given as follows. 1) Current Practice uses the *actual* markups in the holdout period. This reflects the retailer's current decision rules, which are unknown to us. 2) Unit Cost based approach allocates markups proportional to the average unit variable costs of each brand ($\overline{C}_i$) in the last 13 weeks of the estimation data. 3)

Market Share based approach allocates markups proportional to the brands' market shares in the last 13 weeks of the estimation data. This method is counter to common practice in the retail industry, which generally sets lower markups on higher turnover items. For each heuristic as well as our optimization procedure, we compute the posterior mean and standard deviation of the category profit, sales volume, revenue, total variable costs, and unit variable cost in the holdout period ($T = 13$ weeks). The results are reported in Table VI.

[**INSERT TABLE VI ABOUT HERE**]

Table VI indicates that markups derived from the four procedures differ substantially. Looking at each procedure, the variability among brands in our optimal approach appears to be the largest, while that in the current practice seems the lowest. The table also shows that these procedures result in vast differences in profit. Allocating markups proportional to unit costs is only slightly better than the current practice in terms of profits, which are around $110,000 and $108,000, respectively. That difference may well be due to chance. It is clear that allocating markups proportional to the brands' market shares is the best-performing of the three heuristics, and yields a profit around $132,000 in the holdout period. Finally, our optimal markup allocation approach yields the highest profit, around $149,000.

Interestingly, the performance of the market share based markup allocation is closest to the optimal procedure, at least for the data used in this study. This may be of interest to practitioners in the retail industry if the finding can be generalized. Nonetheless, our approach which sets the markups to optimize category profit is clearly superior. A comparison of the optimal markups with current practice suggests that the latter yields a substantially lower spread of the markups. The largest brand (Oral B) is under-priced, while the smallest brand (Aquafresh) is over-priced by the retailer. This explains why setting markups proportional to market shares would increase category profit, and calls into question the common practice in the retail industry of setting lower markups on high turn-over items and higher markups on low turn-over items. This practice may be based on brand-by-brand considerations. But our results suggest that it may be misguided, if one takes the perspective of managing the category as a whole and accounts for interdependencies of sales and prices among brands.

We further explore the causes for the differences in profitability. Table VI indicates that sales levels for the three heuristic procedures are not very different, and the predicted category sales for the optimal markups are only about 6% higher than that for current practice. Revenue for the optimal markups is also only slightly higher than that for the heuristically derived markups, for example, it is only about 1% higher than current practice. But, interestingly, total variable costs (and thus unit variable costs) are much lower in the optimal procedure. For example, unit variable costs are more than 30% lower in the optimal procedure than in the current practice ($0.89 vs. $1.30). These figures reveal that retail profit can be raised without increasing sales volume or revenue. Rather, the higher profit seems to be achieved by setting markups optimally so that (1) a larger proportional of revenue is generated by high markup brands; and (2) (wholesale) prices are driven down by competitive effects. The second point is particularly intriguing: it underscores the importance of including the price component in the model in forecasting future sales and in deriving the optimal category prices. Prior econometric work on category management has often relied on a flexible specification for consumer demand, but has not included a specification for prices. Such an omission can yield an incomplete picture of the main factors contributing to a retailer's profitability and may lead to sub-optimal category pricing decisions.

5 DISCUSSION

Our study meets the needs in the retail industry for sophisticated yet practical decision support tools for implementing category management. The analysis indicates that adopting a model-based CM approach can improve category profit substantially over a retailer's current practice. In addition, the empirical results suggest a simple and effective way for retailers to set markups based on historical sales data by allocating markups proportional to market shares, which could reasonably improve profitability over their current practice. Moreover, we find that larger share brands may be set with higher markups than what is done currently, while the common practice of charging lower markups for high turn-over items may not be in retailers' best interest from a CM perspective. Of course, these findings need to be corroborated in future research before decisive recommendations can be made. Our analysis also reveals that a retailer can achieve higher category profit without an increase in the overall demand, but rather through driving down wholesale prices by competitive force. Interestingly, this empirical finding is consistent with the analytical results of Basuroy, Mantrala and Walters (2001), who also find that CM can increase the category profit for a retailer without increasing total sales of the category. And, more importantly, in accordance with the findings of our study, Basuroy, Mantrala and Walters (2001) show that this profit gain is achieved through lower wholesale prices under the scenarios where a retailer adopts a CM perspective. This convergence of results from two independent streams of research is encouraging.

It has been suggested that the adoption of CM is most effective for categories for which cross-effects are high, and not necessarily for categories with high sales volumes (Zenor 1994). Taking the application of our model one step further, it could be applied to identify the categories where CM is most effective. That would enable the retailer to better coordinate marketing activities across categories. Moreover, a marketer may want to consider a category's strategic role in the overall retail planning.[7] Although profit generation is the primary goal for the vast majority of product categories a retailer carries, other priorities arise in some categories. For example, the so-called "loss leader" categories (i.e., those that are sold nearly at cost) are commonly used to build store traffic rather than to generate profit for those categories themselves. Thus, the objective function needs to be adjusted for decisions on such categories. Although the ultimate goal of category management should be to improve the overall performance of the retailer instead of to maximize the profit of any one category in isolation, we believe that the econometric approach presented in this paper makes a further step in advancing the effectiveness of category management practice.

[7] *Category Management Report*, Grocery Manufacturers, 1995

Table I: Summary Statistics for the Toothbrush Data

Brand	Weekly Sales (units)		Shelf Price ($)		Markups ($)	
	Mean	S. D.	Mean	S. D.	Mean	S. D.
Aquafresh	8.98	7.27	2.77	0.31	1.23	0.06
Butler	14.15	25.65	1.21	0.35	0.87	0.02
Colgate	24.30	15.86	2.64	0.29	1.25	0.04
Crest	16.75	11.37	2.39	0.30	0.96	0.03
Oral B	35.25	24.37	2.54	0.21	0.94	0.02
Pepsodent	10.55	12.09	0.92	0.11	0.47	0.02
Reach	16.70	12.11	2.40	0.26	1.07	0.04

Table II: Posterior Mean and Standard Deviations ($V_\mu^{1/2}$ and $V_\nu^{1/2}$) of Store Constants in the Demand and Price Equations

Brand	Store Constant of log(Sales) Mean	$V_\mu^{1/2}$	Store Constant of log(Price) Mean	$V_\nu^{1/2}$
Aquafresh	1.42[a]	0.232	1.042[a]	0.109
Butler	1.90[a]	0.267	0.584[a]	0.111
Colgate	2.83[a]	0.182	0.916[a]	0.110
Crest	2.46[a]	0.175	0.893[a]	0.110
Oral B	3.32[a]	0.202	0.945[a]	0.109
Pepsodent	1.90[a]	0.236	-0.073[a]	0.110
Reach	2.60[a]	0.230	0.884[a]	0.109

[a] Estimate plus or minus twice the posterior S.E. does not cover the zero value.

Table III: Pooled Demand and Price Coefficients

Equation	Coefficient	Own Price	Cross Price
Sales		$g_0 = -4.35$[a]	$g_c = 0.49$[a]
Price		$a_0 = 0.47$[a]	$a_c = 0.03$

[a] Estimate plus or minus twice the posterior S.E. does not cover the zero value.

Table IV: Posterior Means of Coefficients in the Demand Functions

On\Of:	Aquafresh	Butler	Colgate	Crest	Oral B	Pepsodent	Reach	*Price Index*
Aquafresh	-5.26[a]	0.71[a]	0.09	0.87[a]	0.73[a]	0.37[a]	0.28	-0.06
Butler	0.74[a]	-2.45[a]	0.33	1.96[a]	0.45[a]	0.57[a]	1.67[a]	-0.78[a]
Colgate	0.40[a]	0.31[a]	-4.16[a]	0.35[a]	0.02	0.11	0.97[a]	-0.02
Crest	0.68[a]	0.14	0.75[a]	-4.08[a]	0.02	0.43[a]	0.26	-0.04
Oral B	0.25[a]	0.19[a]	0.63[a]	0.24[a]	-5.78[a]	0.22[a]	0.29[a]	-0.04
Pepsodent	0.01	0.56[a]	0.69[a]	0.08	1.48[a]	-3.51[a]	1.19[a]	-0.97[a]
Reach	0.20	0.15[a]	0.46[a]	0.18	0.65[a]	0.04	-5.22[a]	-0.05

[a] Estimate plus or minus twice the posterior S.E. does not cover the zero value.

Table V: Posterior Means of Coefficients in the Price Functions

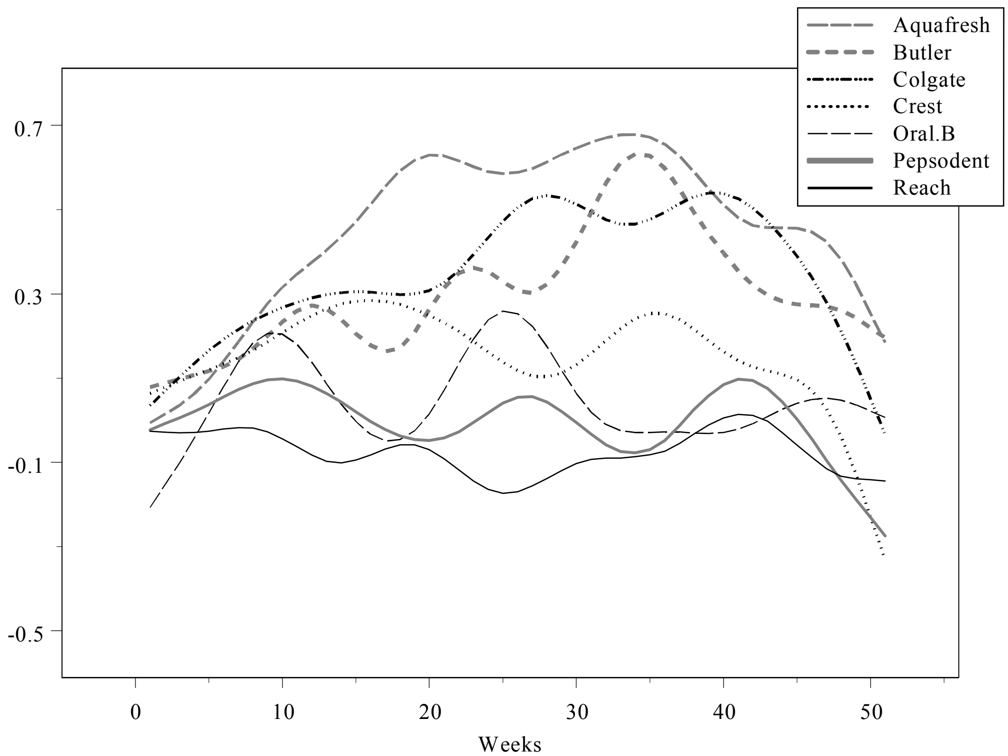
On\Of:	Aquafresh	Butler	Colgate	Crest	Oral B	Pepsodent	Reach	*Sales(t-1)*
Aquafresh	0.45[a]	-0.07[a]	0.14[a]	0.14[a]	-0.05	-0.07[a]	-0.12[a]	0.08
Butler	0.29[a]	0.73[a]	-0.02	0.00	-0.11[a]	-0.01	-0.08[a]	0.59[a]
Colgate	-0.25[a]	-0.01	0.26[a]	-0.06[a]	0.07[a]	0.02[a]	0.27[a]	2.11[a]
Crest	-0.14[a]	-0.06[a]	0.19[a]	0.36[a]	0.43[a]	-0.15[a]	-0.07[a]	1.23[a]
Oral B	0.10[a]	0.01	-0.03[a]	-0.05[a]	0.53[a]	-0.04[a]	-0.13[a]	0.39[a]
Pepsodent	0.01	0.04[a]	0.58[a]	0.24[a]	0.18[a]	0.40[a]	-0.05	0.41[a]
Reach	-0.04[a]	0.03[a]	0.03[a]	0.02[a]	0.09[a]	0.03[a]	0.59[a]	0.82[a]

[a] Estimate plus or minus twice the posterior S.E. does not cover the zero value.

Table VI: 13-Week Markup Allocation and Category Profit Evaluation

Brand / Markups ($)	Current	Unit Cost	Market Share	Optimal
Aquafresh	1.15	1.26	0.38	0.19
Butler	0.86	0.72	0.53	0.63
Colgate	1.24	0.99	1.45	1.20
Crest	0.93	1.08	0.92	0.35
Oral B	0.92	1.24	1.97	2.55
Pepsodent	0.48	0.34	0.63	0.38
Reach	1.10	1.04	0.78	1.39
Gross Profit ($)	107,941.97	110,074.21	131,538.96	148,642.71
SE	3,557.46	3,797.69	5,020.00	8,254.57
Sales (Units)	110,030.60	108,725.66	111,178.02	117,062.01
SE	3,634.79	3,587.63	3,875.27	7,043.64
Revenue ($)	250,549.54	246,402.91	248,731.01	252,957.91
SE	8,563.70	8,400.28	8,632.24	10,216.32
Total Variable Costs ($)	142,607.57	136,328.70	117,192.04	104,315.20
SE	5,058.56	4,621.50	4,236.66	8,285.36
Unit variable Costs ($)	1.30	1.25	1.05	0.89
SE	0.02	0.01	0.01	0.06

Figure I: Smooth Sales-Time Plots for Toothbrush Brands



References

[1] Basuroy S, Mantrala MK, Walters GH. 2001. The Impact of Category Management on Retailer Prices and Performances: Theory and Evidence. *Journal of Marketing* 65: 16-32.

[2] Blattberg RC, George E. 1991. Shrinkage Estimation of Price and Promotional Elasticities: Seemingly Unrelated Equations. *Journal of the American Statistical Association* 86: 304-315.

[3] Boatwright P, McCulloch RE, Rossi PE. 1999. Account-Level Modeling for Trade Promotion: An Application of a Constrained Parameter Hierarchical Model. *Journal of the American Statistical Association* 94: 1063-1073.

[4] Deaton A, Muellbauer J. 1980. An Almost Ideal Demand System. *The American Economic Review* 70: 312-326.

[5] De Crombrugghe D, Palm FC, Urbain JP. 1997. Statistical Demand Functions for Food in the USA and The Netherlands. *Journal of Applied Econometrics* 12: 615-645.

[6] Dorfman JH. 1997. *Bayesian Economics through Numerical Methods*. New York: Springer.

[7] Fan J, Gijbels I. 1996. Local Polynomial Regression and Its Applications. London: Chapman and Hall.

[8] Kloek T, Van Dijk H. 1978. Bayesian Estimation of Simultaneous Equation System Parameters: An Application of Integration by Monte Carlo. *Econometrica* 46: 1-19.

[9] Kopalle PK, Mela CF, Marsh L. 1999. The Dynamic Effect of Discounting on Sales: Empirical Analysis and Normative Pricing Implications. *Marketing Science* 18: 317-332.

[10] Montgomery AL. 1997. Creating Micro-Marketing Pricing Strategies using Supermarket Scanner Data. *Marketing Science* 16: 315-337.

[11] Neslin SA, Powell SG, Schneider Stone L. 1995. The Effects of Retailer and Consumer response on Optimal Manufacturer Advertising and Trade Promotion Strategies. *Management Science* 41: 749-766.

[12] Niederreiter H. 1992. *Random Number generation and Quasi-Monte Carlo Methods*. Philadelphia: Society for Industrial and Applied Mathematics.

[13] Stone JRN. 1954. Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand. *Economics Journal*, 64: 511-527.

[14] Tobin J. 1950. A Statistical Demand Function for Food in the USA. *Journal of the Royal Statistical Society*, Series A, 113: 113-41.

[15] Yang S, Chen Y. Allenby GA. 2003. Bayesian Analysis of Simultaneous Demand and Supply. *Quantitative Marketing and Economics* 1: 251-304.

[16] Zellner A. 1972. *An Introduction to Bayesian Inference in Econometrics*. New York: Wiley.

[17] Zenor, MJ. 1994. The Profits Benefits of Category Management. *Journal of Marketing Research* 31: 202-213.