

Why the Dynamics of Competition Matter for Category Profitability

Sudhir Voleti,

Indian School of Business (ISB), Gachibowli, Hyderabad-50032, India.

Email: sudhir_voleti@isb.edu

Manish Gangwar

Indian School of Business (ISB), Gachibowli, Hyderabad-50032, India.

Email: manish_gangwar@isb.edu

Praveen K. Kopalle,

Associate Dean for the MBA Program, Signal Companies' Professor of Management and

Professor of Marketing

Tuck School of Business at Dartmouth, Dartmouth College, Hanover, NH-03755, USA

Email: kopalle@dartmouth.edu

Why the Dynamics of Competition Matter for Category Profitability

Abstract

Category Management (CM) has become a widespread trade practice in recent years. A category manager's decision problem is complex and multi-faceted owing to demand dependencies across products and across time. Extant research on CM has typically focused on one or the other of these dependencies, but seldom both. The authors address this gap in the extant empirical literature on CM by presenting a competition framework that reconciles cross-sectional breadth (large numbers of SKUs in any given period) with longitudinal depth (demand effects across time). The endeavor is to offer retailers a general, realistic and practical CM approach by comprehensively accounting for competitive effects. The authors demonstrate their approach on real-world data in the beer category for a midsize grocery chain in the US Northeast. Upon determining the optimal weekly prices for the entire assortment over 23 weeks, the authors report a profit yield that is 3.30% more than in the benchmark logit model and substantially more than in the retailer's current EDLP pricing policy.

Keywords: Category Management, Competition Modeling, Pricing, nested Dirichlet Process, Dynamic Programming

INTRODUCTION

Organized retail is an important business sector in mature economies and is growing rapidly in emerging economies. For instance, with \$1.13 trillion of business, retail trade contributed 7.9% to the US economy in 2009 (Retail Trade Federation Report). Category Management (henceforth, CM) has emerged as a popular, rapidly expanding trade practice (e.g., AC Nielsen 2004) in the past two decades. The premise of CM is that category managers will endeavor to co-ordinate marketing decisions (such as on prices, promotions, assortments etc.) across manufacturers to maximize outcomes (e.g., sales, profits etc.) at a category level rather than a brand level. In contrast, the traditional brand centered management role independently manages the marketing mix of individual brands in the category. Almost every major US retailer has adopted CM in some form (Gajanan, Basuroy and Beldona 2007, p.135) and vast majorities of both manufacturers (89%) and retailers (98%) agreed that CM is the most critical issue they face (AC Nielsen 2004). Consequently, the topic of CM has garnered much research and trade attention over the past two decades (Basuroy, Mantrala and Walters 2001; Chintagunta 2002; Gooner, Morgan, and Perreault 2011; AC Nielsen 2004; Zenor 1994).

The retail category manager's problem, of maximizing outcomes over the entire category and not necessarily that of any particular firm, brand or stock keeping unit (SKU), is complex and multi-faceted. Solving the CM problem facilitates answers to several downstream questions of managerial interest. For instance, category managers might want to know which among an Every Day Low Pricing (EDLP), a high-low price promotions strategy, or something in between is best for a particular category at a particular retailer. Which SKUs and brands would do better under EDLP than under a high-low strategy? How long should a promotion be run (e.g., Tsiros and Hardesty 2010 on how best to end promotions)? Which brands and SKUs exert the highest competitive impact on one another and hence can be viewed as close substitutes? How much

incremental gain in outcomes (e.g., sales or profits) measured in dollar terms does the proposed approach deliver as opposed to the benchmark? These questions motivate our work in this paper. We find a systematic way to answer these questions and demonstrate our approach on real-world data.

Extant empirical research that has taken a prescriptive view has largely considered either dependencies across products (following classical competition modeling literature such as Eliashberg and Chatterjee 1985; Sudhir 2001; Wedel and Zhang 2004) or those across time (classical dynamic programming literature starting with Rust 1987). Hence, there exists an opportunity to bring together these two types of dependencies in the context of CM and arrive at an optimal solution. In this paper, we seek to address this gap in the literature. A large body of work on CM is descriptive (e.g., Dhar, Hoch and Kumar 2001; Besanko, Dube and Gupta 2005; Gooner, Morgan and Perrault 2011), analytical (e.g., Raju, Sethuraman and Dhar 2005; Du, Lee and Staelin 2005), cross-category in scope (e.g., Chen et al. 1999), or involves interactions across channel members (e.g., Moorthy 2005; Gruen and Shah 2000). Table 1 displays a subset of empirical work in marketing that either implicitly or explicitly seeks to solve the category maximization problem. The studies are assessed on four dimensions (the four columns) to better annotate the positioning and contribution of this paper.

[Insert Table 1 here]

In sum, we offer retailers a general, realistic and practical CM approach by comprehensively accounting for competitive effects. Our competition framework reconciles cross-sectional breadth (large numbers of SKUs in any given period) with longitudinal depth (demand effects across time). Our approach is prescriptive as well, predicated on a descriptive analysis, in that we compute the optimal set of prices that maximizes category profits over a longer time horizon. We implement our approach on real world data in the beer category from a

mid-size retailer in the US Northeast. We address the retailer's multi-period, category profit maximization problem for a reasonably large number (96) of SKUs. Our results show that our proposed approach yields realistic substitution patterns and cross-SKU effects. Consistent with expectations, we find asymmetric competitive effects, i.e., some SKUs exert greater competitive pressure on others than vice-versa. We demonstrate a case wherein incorporating cross-product demand dependencies (through inter-product similarity and competition variables) in a log-linear framework over a multi-period horizon results in an optimal pricing strategy that is Hi-Lo in nature.

CONCEPTUAL BACKGROUND

Consider a scenario in which a category manager for (say) the beer category wants to set prices for each beer SKU for each week over a specified period (e.g., 6 months after which some SKUs might be added or dropped and the planograms reset) in order to maximize an outcome of interest (such as longer term profitability) for the category as a whole. One approach is to rely on intuition and past experience to identify sets of products with strong demand inter-connections (e.g. Budweiser and Coors at the brand level; cans and bottles of Budweiser at the SKU level, etc.) and thereafter set prices such that total sales are smoothed over and profits are higher than under a less active pricing regime (e.g., markup pricing, competitor benchmarked pricing). Some research suggests that such an approach is not uncommon in that many retailers are unable to deal with the complexity of cross-price effects (McAlister 2005; 2007). While useful and fast when managing a small number of SKUs, this approach rapidly becomes suboptimal when the category hosts a large assortment. A more scientific approach would require that category demand, comprising individual product demands, be formally studied and modeled. Suppose in pursuit of category profit maximization, our focal manager fits a traditional demand model to her sales data, computes own- and cross- price elasticities and develops a picture of how one product

affects another's demand. After all, effective CM requires knowledge of where to allocate scarce marketing resources (shelf space, promotion spends, price discounts and couponing, feature and display, etc.) in order to get the biggest bang for the buck (Dhar et al. 2001). As long as any two products are seen as being substitutes to some extent, one product's characteristics and marketing activities impact the other product's demand. A typical product category in a US grocery store contains about 50 to 100 SKUs and the retail category manager needs to set prices at a weekly level for each of these SKUs while keeping in mind that each SKU's marketing mix affects demand for all the other SKUs. Furthermore, these effects may be asymmetric. Thus, changing Budweiser's 6 pack 12 oz bottle prices by 5% might affect Coors' 6 pack 12 oz can sales much more than vice versa. Compared to the literature available on inter-brand competition (e.g., Wedel and Zhang 2004; Hall, Kopalle and Krishna 2010), there is little research on the intensity and effects of inter-SKU competition, either across brands or within a brand. To determine the competitive nature across all (J) SKUs simultaneously, the category manager needs to estimate an asymmetric (JxJ) matrix of own and cross (or competitive) price effects, which can be quite a challenge due to the number of coefficients involved, many of them with wrong signs etc.

Moreover, even if our focal category manager were to fit a demand model to her sales data, conventional demand models assume independence across time periods (unless dynamic effects are explicitly considered). Such an assumption risks overlooking instances in which the marketing activities of an SKU in one period affect its demand in future periods. For example, forward-looking consumers may either stockpile goods given low current prices or postpone purchases in anticipation of low prices in the next period (e.g., Sun et al. 2003; Kumar and Leone 1988). In the beer example, price promoting Budweiser for a 2-week period not only impacts the sales of Coors in that 2-week period but might also reduce sales of Budweiser (and Coors) in the weeks following the promotion period due to stockpiling and forward buying by consumers. It

may be that the net profit from increased sales at promotional prices (low margins) does not cover the next period's forgone profits from lost sales at full prices (see, e.g., Raju 1992). Other instances from the literature wherein past purchases may influence the probability of future purchases is state dependence in brand choice (e.g., Seetharaman, Ainslie and Chintagunta 1999), or changes in the salience of reference prices to consumers (e.g., Kopalle, Rao and Assuncao 1996). Hence, if demand effects from past periods and/or those on future period sales are not considered, then either optimal actions (from a multi-period perspective) are not taken and money is left on the table, or sales fall below their status quo level as a result of competition effects from past periods. Thus, category managers must also take into consideration the dynamics of product competition – the effect of current prices on future period profitability - and then set each SKU's price by jointly maximizing the category profit. As a result of both these factors, CM in general and category price optimization in particular present a complex problem.

Theoretical Support for Dynamic Effects

A rich body of theoretical work in economics and marketing explains why a single price, EDLP strategy (pure strategy) may not be optimal for firms selling frequently purchased categories (e.g., Varian 1980; Narashimhan 1988; Raju, Srinivasan and Lal 1990; Rao 1991) . Intuitively, firms in competitive markets have an incentive to lower prices to attract price sensitive consumers. Since competitors also follow the same logic, this can lead to an unnecessary price war or suppressed prices for long term. To circumvent this predicament, the firm rather occasionally provides discounts (use mixed strategy) to avoid aggressive competitive response and price wars in the market. This allows a firm, mainly content with the demand from its loyal base who are willing to buy at regular prices, to occasionally gain additional demand from price sensitive segment. In a recent paper, Gangwar, Kumar, and Rao (2014) find in an IRI

Marketing Science dataset that a majority of the SKU's exhibit temporal price dispersion which they argue is a result of inter temporal shifts in demand due to consumer stockpiling.

From a consumer standpoint, one may ask why past prices should affect demand at all, given that lagged price doesn't enter into the consumer's budget constraint or utility directly. One plausible explanation for lagged price effects on current demand is the existence of some type of state dependence (e.g., inventory, reference price, or habit formation). Starting from Guadagni and Little (1983), there has been considerable effort and a rich body of literature in marketing devoted to understand the various types of state dependence.

Factors Affecting Optimal Prices

To demonstrate our profitability analysis, consider SKU prices over time, which in retailing have a strong impact on customer behavior and retailer profitability (Ma, Ailawadi, Gauri, and Grewal 2011; Levy, Grewal, Kopalle, Hess 2004). To price SKUs optimally, retailers need to consider at least four factors in their pricing decisions: (1) Product Attributes, which are essentially SKU characteristics such as brand, size, type, etc. (2) Price sensitivity, i.e. change in an SKU's demand with changes in its price. (3) Substitution or competitive effects, i.e., the change in an SKU's demand changes due to changes in other SKUs' characteristics (both measured and latent) and marketing mix. The retailer would thus need to evaluate the relative demand effects and differential SKU margins before changing the price of any focal SKU. (4) Dynamic effects of marketing mix, which comprise the competitive effects on demand over time (on both self and on other SKUs). For instance, would changing an SKU's price today affect its own demand tomorrow? Would it affect other SKUs' demands in future periods? We note that these factors are not limited to consumer packaged goods in brick-and-mortar stores, but extend also to electronic channels as well as to information products (e.g., Kannan, Pope, and Jain 2009).

One of the difficulties in producing a useful demand model that appropriately captures all the competitive effects is the sheer number of SKUs within a category. Prior research has typically captured inter-product competition via variants of the logit model (Ben-Akiva and Lerman 1989; Guadagni and Little (1983); Sriram, Balachander, and Kalwani 2007) that model the probability of purchasing a product. The logit formulation of demand relates purchase probability (or market share) to a set of covariates (that includes the marketing mix of all the products and their respective characteristics) in the denominator, thereby implicitly accounting for competition effects. Hence, the logit serves as a logical baseline against which candidate approaches to incorporating competition can be evaluated. In this regard, although the BLP framework (Berry, Levinsohn and Pakes 1995; Sudhir 2001) can incorporate both heterogeneity and competitive effects, it (i) does not consider all SKU attributes in the estimation (due to collinearity across attributes), (ii) does not explicitly account for similarities across SKUs (either in observed or in latent attributes), and (iii) faces estimation challenges in the presence of a large number of SKUs or the inclusion of dynamic effects. In this paper, our approach addresses these shortcomings and provides a more profitable pricing path for the various brands in the category.

One paper that is similar in spirit to ours is Shah, Kumar and Zhao (2014), which takes a CM perspective in modeling assortment decisions among heterogeneous retailers who do not set prices, operate in a limited SKU setting and in an underpenetrated (i.e. potentially non-stationary) market in the developing world. Our work complements Shah et al. (2014) in that ours is relevant in a stable, large fixed assortment context where we study aggregate demand impact of dynamic SKU pricing decisions. We focus on a single retailer who centralizes price-setting in a developed market. In this context, we consider a simple but intuitive way of capturing competition that is dynamic in nature and is based on inter-SKU similarities. We then conduct our estimation and optimization accordingly to arrive at prescriptive results.

A FRAMEWORK FOR COMPETITION

Recall our previous illustrative example wherein the beer category manager, in order to maximize category profits, needs to resolve two types of complexities - dependence across products and that across time. Competition lies at the core of both these types of demand dependencies. A lot of research attention has focused on analyzing competition to disentangle its components (for e.g., see Moorthy 2005). Two products can have innate similarity in fixed attributes (e.g., Miller Lite and Bud Light share the 'light beer' attribute). Alternatively, perceived similarity between two products can be affected by changes in their Marketing mix (e.g., 'Budweiser Select' is priced much higher than regular Budweiser or Miller and is typically not seen as competing with them). In what follows, we link fixed attributes of competition (both in the space of observed data and latent parameters) and the time-varying marketing mix attributes into a unified framework of competition that accommodates (i) both the inter-SKU and the dynamic aspects of competition, (ii) the category's brand-SKU hierarchy, and (iii) a mechanism to facilitate price optimization downstream.

In their work on time-varying brand equity, Voleti and Ghosh (2013) consider static inter-SKU competition for linear demand models. However, their model has limitations. First, although it accounts for cross-product competition, it rules out dynamic effects, and therefore own-product competition. This is limiting. Second, their optimal price paths would imply an independent application of the inverse elasticity rule in each period for each product, and thus restrictive. We extend their general framework by dynamically connecting changes in the Marketing Mix (henceforth MMIX) of any product to changes in demand both across SKUs and across time.

Let $COMPTN(jt, it)$ be the competitive impact on the sales in period t of focal SKU j due to rival SKU i . At least two main factors impact $COMPTN(jt, it)$: inter-SKU attribute similarity

and marketing mix effects. Let $SIMIL(i, j)$ denote the degree of attribute similarity between i and j , and $MMIX_{it}$, the endogeneity corrected marketing mix activity of SKU i in period t ¹. While MMIX effects can be contemporaneous (e.g., Eliashberg and Chatterjee 1985; Kumar et al. 2009; Kannan and Yim 2001), consumer behavior such as stockpiling, consumption, and purchase postponement (e.g., Sun, Neslin and Srinivasan 2003) may result in inter-temporal MMIX effects. Let k periods of MMIX lags, denoted by $MMIX_{i,t-1}$, $MMIX_{i,t-2}$, ... $MMIX_{i,t-k}$ impact j 's demand. Let the term $MMIX.EFFECT_{i,t,k}$ denote the combined effect of the rival product i 's marketing mix activity in period t from 0 to k lagged periods. We model a multiplicative relation between $SIMIL(i, j)$ and the $MMIX.EFFECT$, effectively implying that the degree of similarity between i and j scales up or down the MMIX effect of i on j . Thus:

$$COMPTN(jt, it) = SIMIL(i, j) * MMIX.EFFECT(i, t, k). \quad (1)$$

Since $SIMIL$ represents the degree of attribute similarity between products, it can be inversely related to an inter-product 'distance' in attribute space (e.g., Carpenter and Nakamoto 1989 for preference structures). Let $l = 1, 2, \dots, L$ represent observed, discrete attributes of products $j=1, 2, \dots, J$ at the SKU level. Let indicator function $I(.)$ assign a product-pair (i, j) a value of 1 if it shares attribute l , and 0 otherwise. Since not all product attributes are created equal in their impact on product similarity, we use a set of parameters to differentially weigh attribute importance in contributing to competitive sales effects. Assuming a linear function for simplicity, $SIMIL(i, j)$ is:

$$SIMIL_{ij} = \sum_{l=1}^L \delta_l * I(l_i = l_j). \quad (2)$$

¹ Some variables such as the time-varying MMIX elements may be strategically set by manufacturers and retailers. We treat these potentially endogenous variables using a set of instruments comprising exogenous or pre-determined variables (detailed in the data section) to obtain a set of endogeneity corrected values MMIX.

The attribute importance weights are assumed to be proportional to the marginal sales impact of the corresponding attributes drawn from an auxiliary analysis sample. Voleti, Kopalle and Ghosh (2015), in their static two-stage demand model based on a nested Dirichlet Process, extend the COMPTN variable in equation (1) to include as arguments the effect of unobserved attributes in latent parameter space. We propose to extend the COMPTN approach to include the effects of current and past MMIX activity as well. We provide a detailed description of the development of SIMIL as well as other aspects of competition in Supplementary Appendix A. $MMIX.EFFECT_{i,t,k}$ captures i 's current and past MMIX activity as a general function of $MMIX_{it}, \dots, MMIX_{i,t-k}$:

$$MMIX.EFFECT_{i,t,k} = \beta_0 Ln(MMIX_{it}) + \dots + \beta_k Ln(MMIX_{i,t-k}). \quad (3)$$

Per Equations (1)-(3), the term $COMPTN(jt, it)$ in Equation (1) captures the competitive impact - both contemporaneous and lagged upto k periods - of rival SKU i on focal SKU j in period t . We now aggregate $COMPTN(jt, it)$ to arrive at the overall competitive impact due to all rival SKUs $i \neq j$ on focal SKU j in period t :

$$COMPTN_{jt} = \sum_{i \neq j} COMPTN(jt, it). \quad (4)$$

Thus, we incorporate competition in a dynamic fashion via inter-SKU attribute similarity and past and current MMIX levels and yields, for every SKU j in every period t , the net competitive impact on j 's sales due to the MMIX actions of all other SKUs in time t to $(t-k)$. Further, consistent with findings in the marketing literature (e.g., Russell and Kamakura 1994), we find the inter-product competition is asymmetric. That is:

$$COMPTN(jt, it) \neq COMPTN(it, jt) \quad (5)$$

unless two similar SKUs have identical contemporaneous and lagged MMIX levels.

Overall, the COMPTN variable as constructed primarily provides a mechanism to transmit the effects of changes in any product's marketing mix to all other products' sales in both current and future periods.

Demand Specification

We start with a basic specification of the popular log-log demand model (e.g., Hoch et al. 1995; Wedel and Zhang 2004) in which the log of sales quantity is regressed over a set of demand determinants. The logarithmic function offers a natural diminishing-returns pattern, accommodation of various response shapes and rates (e.g., Lilien, Kotler and Moorthy 1992) and a ready interpretation of coefficients as elasticities. Let $j = 1, \dots, J$ index SKU and $t = 1, \dots, T$ index time-period. Thus:

$$\begin{aligned}
 \text{Ln}(SALES_{jt}) &= \alpha_0 + \mathbf{Z}_t \boldsymbol{\delta}_1 + \text{Ln}(\mathbf{MMIX}_{jt}) \boldsymbol{\alpha}_{1,j} + \text{Ln}(\mathbf{MMIX}_{j,t-1}) \boldsymbol{\alpha}_2 + \dots \\
 &\quad + \text{Ln}(\mathbf{MMIX}_{j,t-k}) \boldsymbol{\alpha}_{2+k} + \text{COMPTN}_{jt} \alpha_{3+k,j} + \alpha_{4+k} * \text{SIMIL}_{j,t} + \eta_j + \varepsilon_{jt}, \\
 \varepsilon_{jt} &\sim \text{IID } N(0, \sigma_\varepsilon^2), \quad j = 1, 2, \dots, J; \quad t = 1, 2, \dots, T, \\
 \boldsymbol{\theta}_j &= [\boldsymbol{\alpha}_{1,j} \quad \alpha_{3+k,j} \quad \eta_j] \sim F(\cdot).
 \end{aligned} \tag{6}$$

Here, \mathbf{Z}_t is a vector of control variables (e.g., seasonality), log is taken for continuous variables, and $\boldsymbol{\delta}_1$ is the corresponding coefficient vector. $\text{Ln}(\mathbf{MMIX}_{jt})$ is a vector of SKU-specific log price, log promotion (or advertising) and log distribution variables in period t corrected for endogeneity and $\boldsymbol{\alpha}_{1,j}$, a vector of MMIX elasticities specific to SKU j . $\text{Ln}(\mathbf{MMIX}_{j,t-k})$ represents the k -period lagged MMIX effects and $\boldsymbol{\alpha}_{2+k}$, the corresponding MMIX elasticities. Parameters, $\boldsymbol{\alpha}_{2+k}$, are not SKU-specific to avoid making excessive demands on the data and to preserve degrees of freedom. This formulation is general and is based on knowing the lag-length ' k '. Further, since the log-log model corresponds to a multiplicative model, we log-transform the variables 'Distribution' (DISTBN) and 'Promotion' (PROMO), which may take zero values, using

the transformation $\log\text{DISTBN} = \text{Ln}(1+\text{DISTBN})$. This avoids computational traps. We did a robustness check regarding our use of the constant '1.0' with other values (0.1, 0.01 etc) and found qualitatively similar model results. COMPTN_{jt} summarizes the dynamic effects of competition on SKU j due to all relevant SKUs in period t and $\alpha_{3+k,j}$, the average sales impact of competition on SKU j . Similarly, $\text{SIMIL}_{j\cdot}$ represents the average extent of attributes shared between SKU j and the rest of the assortment, weighted by attribute importance with (homogeneous) coefficient α_{3+k} (since SIMIL doesn't vary with time). Please note that both COMPTN and SIMIL variables in Equation (6) have been constructed prior to estimating the demand model using parameter weights from an auxiliary dataset (which does not feature in the subsequent analyses). The aggregate sales impact of all time-invariant product attributes is captured directly in the SKU-specific random effects term, η_j . The measurement error ε_{jt} is drawn IID from a mean zero normal distribution.

One implication of the model in Equation (6) is that the competitors' MMIX has long term effects on future focal brand sales (SALES_{jt}). We note that there are primarily two possible ways in which a competitor's (say, i) MMIX in past periods (say, $t-k$ for $k>0$; let $k=1$ for a one period lag) can affect focal SKU's (say, j) current sales SALES_{jt} . One, a direct effect where SALES_{jt} is a direct function of $\text{MMIX}_{i,t-1}$, and two, an indirect effect whereby $\text{MMIX}_{i,t-k}$ affects $\text{MMIX}_{i,t}$ which in turn affects SALES_{jt} . There is no strong theoretical or intuitive support for the first path in the absence of some type of state dependence. In contrast, the second path is clearly a state dependent approach but raises the question of how to model SALES_{jt} without explicitly bringing in MMIX_i into the RHS of Equation (6) (owing to model parsimony considerations). This is where the COMPTN term as formulated helps. We model SALES_{jt} as a function of

\mathbf{MMIX}_{jt} , $\mathbf{MMIX}_{j,t-1}$ and COMPTN_{jt} where COMPTN_{jt} summarizes the indirect sales effects on SALES_{jt} of all $\mathbf{MMIX}_{i,t}$ & $\mathbf{MMIX}_{i,t-1}$.

The set of SKU-specific random parameters $\left[\boldsymbol{\alpha}_{2,j} \quad \alpha_{3,j} \quad \eta_j \right]$, denoted by θ_j , are drawn from a general distribution $F(\cdot)$. Different specifications for $F(\cdot)$ yield different models for treating parameter heterogeneity. For instance, in the traditional approach, $F(\cdot)$ is a multivariate Gaussian hyperdensity, $\boldsymbol{\theta}_j \sim N(\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_\theta)$ which yields the normal-parametric (or N-P) model (Conley et al., 2008). However, for flexibility, robustness against particular distributional assumptions as well as advantages that accrue to the Bayesian estimation of complex models (e.g., exact finite sample inference, data augmentation for latent quantities), we implement a Bayesian Semiparametric approach based on a Dirichlet Process prior (e.g., Antoniak 1974). We express the Dirichlet Process Prior (DPP) based model, parameterized by a concentration parameter α and a base distribution \mathbf{H} as:

$$\boldsymbol{\theta}_j \sim DP(\alpha, \mathbf{H}), \quad j = 1, 2, \dots, J. \quad (7)$$

Note that our unit of analysis consists of two distinct product levels, namely, brand and SKU, which have a clear hierarchical relationship in that brands deploy SKUs. Modeling inter-related product demand would require that similarities and dissimilarities at these two distinct product levels be accommodated. The importance of doing so has been recognized in the literature (e.g., Fader and Hardie 1996; Wedel and Zhang 2004). Brand hierarchies imply a tiered structure which reflects firms' branding strategies (e.g., Rao, Agarwal and Dalhoff 2004) and have implications for category structure (e.g., Fader and Lodish 1990), and at a more abstract level, the market structure (e.g., Erdem 1996). Both manufacturers and retailers work with product lines which have typically brand hierarchies embedded in them. In the unlikely scenario that demand dependencies across products are known a-priori, the nested logit model (Kannan

and Wright 1991; Sudhir 2001) can account for a tiered product assortment. Kumar et al. (2009) use a weighted random coefficient model to explicitly model competition across adjacent price and brand tiers. We take the approach of Voleti et al. (2015) who use a nested Dirichlet process (Rodriguez et al. 2008) to uncover demand dependencies with minimal a priori assumptions about competitive structure.

To justify their use of a more complex demand model (the nested Dirichlet Process or nDP based approach) as against a simpler alternative (the normal parametric), Voleti et al. (2015) perform a model assessment and validation exercise. We replicated their exercise and found that the nDP significantly outperforms the normal parametric on a number of measures including model fit, predictive ability and accommodation of category branding constraints (see Supplementary Appendix B). More generally, our modeling approach accounts for brand influences that act through SKUs in at least 3 ways. First, the nested Dirichlet process (nDP) demand specification explicitly acknowledges and accounts for the brand-SKU hierarchy by nesting similar-SKU groups ('child' level clustering) within a 'parent' level clustering of similar-brands. The groupings here are endogeneously derived and represent an improvement over those supplied ex ante in the nested logit case. Second, in our competition submodel, brand membership of SKUs is among the time-invariant product attributes considered that make up $SIMIL(i, j)$ which is one of two arguments required to specify the $COMPTN(it, jt)$ term. And Third, in our actual optimization routine, we optimize the brand level price vectors directly (described in the following section).

ENHANCING CATEGORY PROFITABILITY

Given demand model parameters, improving profitability would now require a comprehensive price optimization exercise. The issue is inherently dynamic because of inter-temporal effects, i.e., price changes from past periods affecting current period demand. We next

illustrate how the inter-related product demand helps improve category profitability in the context of a retailer's multi-period, profit maximization problem.

Maximizing Profitability

The retailer maximizes category profit, i.e., maximizes the sum of discounted profits over a finite time horizon $t = 1, 2, \dots, T$ by selecting a sequence of SKU prices over time, thus:

$$\text{Max}_{\{PRICE_{jt}\}_{t=1}^T} \sum_{t=1}^T \sum_{j=1}^J \left[\Theta^t (PRICE_{jt} - COST_{jt}) * \hat{S}_{jt} \right] \quad (8)$$

Here, Θ is the discount rate, \hat{S}_{jt} is the expected sales for SKU j in period t using the revised demand model specified in Equation (6), and incorporating the dynamic, competition variable $COMPTN^{(nDP)}$, the marketing mix (comprises price, promotion and distribution variables in our empirical application) of SKU j in period t .

The retailer's decision problem can now be reformulated as follows. Given the initial values (at time, $t=0$) of the marketing mix variables, select SKU prices in each period that maximizes the objective function in equation (8). The retailer's category profit maximization problem is thus formulated as an optimal control problem in discrete time, with past marketing mix variables as the state variables and current period prices of the SKUs as the control variables. We use dynamic programming to solve this. We provide a more detailed description of the normative dynamic programming model in Supplementary Appendix C. Please note that we use expected sales in our optimization model. In theory, the error in sales (Equation 6) should be a state variable. However, our model is already cursed with the dimensionality problem and making the error a state variable will only exacerbate the issue. Prior approaches on price optimization has relied on expected sales (Ailawadi, Kopalle, and Neslin 2005; Basuroy, Mantrala, and Walters 2001) and we follow this literature in the paper.

Benchmark Model

The traditional multinomial logit formulation implicitly accounts for competitive responses by including rival products' MMIX variables in its denominator. As such, it is an appropriate benchmark model both for the efficacy of the competition function in aiding demand estimation and for dynamic price optimization. To maintain comparability with the proposed linear demand specification, we model a mixed logit demand model with SKU specific parameter heterogeneity for fixed attributes and price. The lagged MMIX variables however are modeled as homogeneous parameters, in keeping with the proposed linear demand specification. We find that the most important MMIX variable – price per unit volume (or size normalized price) - shows limited intra-brand variation across SKUs, for every brand in the sample, relative to inter-brand variation. Hence, we use normalized price at the brand level as our price variable in all the models we estimate. This enables an apples-to-apples comparison across models. Let $j=1,2,\dots,J$ index SKUs in the sample. Let MMIX_{jt} and $\text{MMIX}_{j,t-1}$ be contemporaneous and lagged marketing mix vectors for SKU j in time t , and let the corresponding coefficients β_j and γ_k respectively. Let α_j denote the random intercept for SKU j , and s_{jt} denote its sales share in time t . Then:

$$s_{jt} = \frac{\exp(\alpha_j + \mathbf{Z}_t \delta_1 + \beta_j \text{MMIX}_{j,t} + \gamma_1 \text{MMIX}_{j,t-1} + \dots \gamma_k \text{MMIX}_{j,t-k})}{1 + \sum_{a=1}^J \exp(\alpha_a + \mathbf{Z}_t \delta_1 + \beta_a \text{MMIX}_{a,t} + \gamma_1 \text{MMIX}_{a,t-1} + \dots \gamma_k \text{MMIX}_{a,t-k})}. \quad (9)$$

The benchmark mixed logit model in Equation (9) connects price changes made in the course of the backward induction algorithm to own and cross-product sales impact via the contemporaneous and lagged marketing mix variables. In the logit model, to arrive at predicted sales we need an estimate of the total market size, M_t , for the product category at the focal retailer. Previous studies that have used the logit (e.g., K. Sudhir 2001) do so in different ways. We fixed the market size to ten times the maximum category sales recorded in the dataset. We

tested for the sensitivity of estimated mixed logit parameters to changes in the multiple (pegged alternatively at 20 and 50), and found qualitatively similar results. Thus, the predicted sales figure is given by:

$$\hat{S}_{jt} = PRED.SALES_{jt} = \hat{s}_{jt} * M_t. \quad (10)$$

Figure 1 depicts a flowchart of the various model components and the estimation steps in our proposed approach.

[Insert Figure 1 here]

This completes our model specification. In the next section, we implement the model on real world data in the beer category and analyze the results.

EMPIRICAL APPLICATION

Data Description

Since our approach takes the perspective of the retail category manager's profit maximization problem, for relevance and consistency, we use data from a single retail chain. We take beer category data from 56 stores of a midsize grocery chain in the US Northeast. This retailer has centralized pricing and assortment planning in all its stores. We have 23 weeks of sales and marketing mix data for 96 SKUs from 15 brands yielding a total of 2171 usable observations. The time-varying MMIX variables available are DISTRIBUTION (% stores in which the SKU was available in a particular week), PROMOTION (% SKU sales made on any promotion in that week), PRICE per fluid ounce of product for that SKU in that week and unit COST per fluid ounce (as price instrument) for that SKU in that week. The national ADSPEND of each brand for that year is also available and is modeled as a fixed effect because only cross-sectional variation across brands for this variable is available in our data. Control variables were MONTH dummies (Jan to May) for seasonality control. The time invariant product attributes available in the data are BRAND, CONTAINER type (bottle/can), beer TYPE (ale, light, craft,

regular), beer COLOR (light, amber, dark, golden), PACKAGING types (6P 12oz, 12P 12 Oz, 18 Pack and 24 pack). We use this fixed product attribute information in two ways - as price instruments in a hedonic regression and as an input to computing SIMIL - the product attribute based inter-SKU similarity measure.

[Insert Table 2 here]

Table 2 summarizes the variables used in the analysis. Considerable variation is seen in both the dependent and the MMIX variables. The instrumental variables for the three time-varying MMIX elements consist of exogenous and upstream variables (season dummies; Bureau of Labor Standards data on Purchasers' Price Indices for material inputs to beer, namely malt and barely, as well as Aluminum--all used as price instruments; regional food inflation indices from the Bureau of Labor Statistics) in addition to the fixed product attributes. The R-square of the instrumental variables regressions are well above 85% and the respective correlations between the actual and predicted MMIX variables are also high (above 0.75).

Model Details

In the interest of brevity, we move the descriptive detail of the choice of priors for the Bayesian model parameters, the modeling choices, convergence diagnostics and related robustness checks to Supplementary Appendix D. Further, in supplementary Appendix E, we provide the code used to estimate the linear demand model. Below, we elaborate on the choice of lag length and the form of the price input to the dynamic programming algorithm.

To find the appropriate lag length k for marketing mix effects in Equation (3), we varied k from 1 to 4 and computed the competition variable for the sample each time. We found that the competition terms for $k=1$ to 4 are highly correlated (well above the 0.90 level). Hence, for simplicity and tractability, we use $k=1$ as the appropriate lag length. Likewise, we found that

coefficients of lagged own-MMIX elasticities in a mixed effects regression rapidly decayed for $k > 1$. Hence, we chose $k=1$ in Equation (6).

There may be environmental and institutional constraints that we as researchers may not observe directly in the data. For instance, prices above certain latent thresholds may induce consumers to switch retailers entirely. Hence, to ensure the results are within reasonable bounds, in determining the profitability, we constrained the prescribed prices to remain within the range of the currently set prices. Similarly, we constrained unit profit margins to not vary far (over 1.5 standard deviations) beyond the observed range of profit margins. This helps rule out instances of unrealistic model predictions. We set the normalized prices of all SKUs within a brand at the mean value for the brand for ease of computation and tractability in the dynamic programming stage. Since the price per fluid ounce changes very little across SKUs within a brand, we expect minimal impact. By the interpolation of the state space, we expect to obtain prescribed policy paths of the prices of the 15 brands in the sample over 23 weeks.

We next discuss the results of our analysis and the corresponding managerial insights.

RESULTS

Competition Results

First, we summarize the competition variable. The mean and standard deviation of non-normalized COMPTN is 444 and 205 respectively. For better interpretability, in the demand model, we normalized COMPTN to have a mean of one. Likewise, the average SIMIL across all SKU pairs is normalized to 1. The primary advantages of using this term are that it is intuitive, easy to compute cross-product marginal effects, and substitution patterns are implied. To demonstrate this, we present an example in Table 3. We chose three SKUs having high, medium,

and low SIMIL vis-a-vis one another for illustrative purposes. The aim is to show (in a 3x3 table) how cross-price elasticity, SIMIL and COMPTN vary across them.

- SKU #56 is the largest selling SKU in our data for the focal retail chain: A light colored Canadian import called 'Labatt Blue Light' 12Pk 11.5 Oz Bottle (priced in the low tier).
- SKU #49 reports the highest attribute similarity with SKU #56. It is the Labatt Blue Light 6 Pk 11.5 Oz Bottle.
- SKU #42, which shows the lowest SIMIL score with SKU #56, is Heineken 6Pk 12 Oz Can - a golden colored regular (i.e., not light or low-calorie) brew (priced in the super premium tier).

Table 3 shows for each focal SKU (in the rows), the cross-price-elasticity with the rival SKU (in the columns), the SIMIL(focal, rival) score in square brackets [], and the average COMPTN impact of the rival on the focal SKU in curly brackets {}.

[Insert Table 3 here]

The diagonal cells show only the own-price elasticities since SIMIL and COMPTN for an SKU with itself is not defined. Among the off-diagonal cells, expectedly, the cross-price elasticities are positive implying that a rise in the price of substitutes causes focal SKU sales to rise. The magnitudes indicate that substitution effects correlate strongly with inter-product similarity. Other interesting details emerge. For instance, (i) attribute similarity between SKUs #56 and #49 is very high [SIMIL = 1.8660 against an average of 1] whereas that between #56 and #42 is at the minimum possible value [SIMIL = 0]. (ii) The inter-product COMPTN effects demonstrate asymmetry. E.g., the competitive effect of SKU #49 on SKU #56 is 9.711 whereas that of #56 on #42 is 10.4977. Since SIMIL between dissimilar SKU pairs is low, so is the COMPTN effect. Thus for example, SIMIL between #56 and #42 is zero and hence so are both COMPTN and cross-price elasticity. Similarly, SIMIL between #42 and #49 is very small (0.08

when the average is 1), and consequently, so are the corresponding COMPTN and cross-price elasticity values.

Demand Model Results

One may ask whether and to what extent the proposed additions to the log-linear demand specification, namely, SIMIL and COMPTN affects model performance. We assess the value of including the SIMIL and COMPTN terms in the linear demand model by estimating four models as part of a 2x2 grid - with and without SIMIL, against with and without COMPTN - and comparing the complexity penalized fit in the holdout prediction sample across them.

[Insert Table 4 and Figure 2 here]

Table 4 displays these results. We find that the 'full' model specification - having both SIMIL and COMPTN - yields the best predictive fit to data compared to alternative specifications. Even on calibration sample fit, the proposed model specification does best. Henceforth, for results and discussions, we use the full model specification (both SIMIL and COMPTN) as our proposed demand model.

To assess the relative importance of the constructed measures of inter-product similarity and competition, following Silber et al (1995), we compute the contribution to explaining variance in sales of different groups of regressors, namely, [Contemporaneous own] Price effect on sales, [Own] Lagged Price effect on sales, non-price MMIX effects, cross-SKU sales effects [measured by SIMIL and COMPTN], [Other independent and] Control variables. The results are displayed in Figure 2. We find that although variation in products' own-price dominates variance explained in sales (close to two-thirds of the modeled effects), lagged prices play an important part (explaining ~18% of the modeled variation in Y) emphasizing the need to account for dynamics. This is followed by non-price MMIX variables (distribution and promotion, in our

data) at ~11%. Finally, at over 5% of the modeled variance in sales, the need to account for cross-product competitive effects is brought out.

Table 5 displays summaries of the posterior draws of the full log-linear demand specification's main effects - both from SKU-specific random parameters (Panel A) and from homogeneous parameters (Panel B). In Panel A, the parameter estimates appear to bear face validity. For instance, we find that the price parameter is negative in sign whereas the promotion and distribution parameters are positive. We find that, as expected, the marginal effect of competitive intensity on sales is negative. The demand model's main strength in the present application context is that it parsimoniously captures the dynamic impact of competition and at the same time accommodates category branding structures.

[Insert Tables 5 and 6 here]

In Table 5's panel B, we find that advertising spend positively correlates with sales; that high prices for a product depress demand not just in the current period but also in the next time period (lagged price coefficient is negative and significant)²; and that the distributional reach or promotional intensity of a product in the past week has no significant effect on sales in the current period. We find that a product's overall SIMIL score correlates positively with sales. Recall that the variable SIMIL for each (focal) SKU is simply the average number of shared attributes with every other SKU weighted by attribute importance. Thus, SIMIL's operationalization in this context can be interpreted as a measure of attribute popularity. For instance, *Ceterus paribus*, bottles sell better than cans and light beer outsells regular beer (based on auxiliary sample regression results). Hence, high-selling SKUs are likely share these 'popular'

² The intuition, following Gangwar et al (2014) is that deep discounts in one period cause loyals to stockpile leaving only weak-preference switchers in the market in the next period, to attract whom the firm has an incentive to discount again in the subsequent period as well. Hence the correlation in prices over consecutive periods.

attributes with one another. In turn, this implies that a high SIMIL score may correlate strongly with sales.

Finally, Table 6 displays the benchmark model (mixed logit) results. The logit model's parameters do not have a 'direct' or marginal effects interpretation. However, the sign of the effects remain indicative of the direction of impact. Thus, we find that the price parameter exerts a negative influence on market share (aggregation of purchase probabilities) whereas that of promotion or distribution is positive on the average. To compare model fit between the logit and the competition based models, we calculated the root mean square error based on the prediction error in $\log(\text{sales})$. We find that both models seem comparable in terms of fit in both the estimation and the holdout samples. Thus both the logit and the linear models appear to approximate the true demand model in a manner consistent with one another's fit and results. However, relative model performance in the optimal pricing algorithm remains to be seen. We used the parameter estimates both from the log-linear model and the mixed logit models as inputs to the backward induction, profit enhancing algorithm, results of which are discussed next.

In sum, we find that accounting for both cross-sectional and longitudinal competition improves outcomes. In the demand model, both model fit and prediction in a holdout sample are shown to be improved with SIMIL and COMPTN than without them. Further, we find that SIMIL and COMPTN together explain 5.5% of variation in sales - the third most impactful group of regressors after Price and non-price marketing mix variables. Introducing the dynamics of competition also hugely impacts outcomes. We find that Lagged price as a determinant of demand alone explains 18.2% of the variation in sales.

Optimal Price Paths

The backward-induction based dynamic programming algorithm provides a set of category-profit maximizing 'optimal' prices for each brand in each time period. Figures 2A, 2B

and 2C show the observed price paths, the optimal price paths under the log-linear model and that under the mixed logit specification respectively for five of the best known and highest selling brands at the retailer, namely - Labatt, Budweiser, Miller, Corona and Heineken. The y-axes are the same scale in both plots to enable better comparison.

[Insert Figures 3A-3C here]

We note the following five salient points. First, the price paths under the current pricing policy (Figure 3A) seem to follow an EDLP policy - stable over several weeks followed by changes in small increments. The average percent difference between the highest and lowest prices for each brand in the sample period is 4.4% under the current price path. The corresponding figures for the linear and logit models are 27% and 11% respectively.

Second, the log-linear model (Figure 3B) prescribes some variation in pricing levels across weeks and resembles a Hi-Lo pattern more than an EDLP one. Based on prior literature in marketing which suggests that High-low pricing policies are generally more profitable relative to a constant price strategy (e.g., Hoch, Dreze and Purk 1994; Mazumdar et al. 2005; Kopalle et al. 1996), ex ante, our category profit expectations would be higher from a hi-lo pricing path than from an EDLP one.

Third, the logit model prescribes a price path that resembles an EDLP policy in that price volatility appears to be relatively low. Hence, ex ante, we would expect the logit's category profit to be closer to the current price path (and hence, lower than under the linear price path).

At this stage one may ask why the proposed and baseline models differ in their optimal strategies. The proposed log-linear, COMPTN model and the baseline mixed logit model both account for inter-product demand dependencies or competition effects, but do so in different ways. The logit approach models the focal SKU's share of sales as a function of the ratio of the demand effects of the focal SKU's attributes to the aggregate demand effect of the attributes of

all the SKUs in the sample. This latter quantity - the aggregate demand effect of the all the SKUs' attributes - appears in the denominator in the RHS term in the model. Note that even though parameter heterogeneity for attribute effects is allowed in the mixed logit model, this denominator term is identical across SKUs (E.g., taking logs on both sides of the logit model yields the denominator as the inclusive value - a constant term - on the RHS).

In contrast, the log-linear model constructs a COMPTN term to capture inter-SKU demand dependencies as a variable of interest in the dataset. Further, the effect of this COMPTN term on focal SKU sales is modeled as heterogeneous parameters (under a robust, flexible nested Dirichlet density). Thus, the proposed model is able to account for SKU-specific heterogeneous competition effects whereas the mixed logit model assumes a homogeneous competition effect across SKUs. We include the exploitation of this property of the proposed model as part of the paper's incremental contribution. Due to the difference in the way competition effects are captured in the two model specifications, we see very different optimal outcomes - Hi-Lo versus EDLP, i.e., the corresponding managerial implications emerge.

Fourth, to better assess the magnitude of the similarities and differences between the prescribed price paths and the current one, we analyze the statistical correlation between the current and the prescribed price paths as well as the rank correlation between brands based on the average current and prescribed prices. Table 7 shows the correlation coefficients. The lower triangular cells contain the Pearson correlation between price paths and the upper triangular cells (shaded region) contains the Spearman rank correlation coefficients. Table 7 indicates that although the logit model prescribes a price path that is correlated with the current one, it changes the order of the brands within price-tiers much more than the proposed COMPTN model does. The log-linear, COMPTN model thus appears to yield a realistic optimal price path which while varies, does so within the same price band as the product's current pricing. E.g., Labatt is the

lowest priced brand and Corona the highest priced one in both Figures 3A and 3B. However, the order changes in Figure 3C (logit) where Heineken moves to the top slot.

[Insert Table 7 here]

Fifth, Table 7's last column shows the mean and standard deviation of the observed prices under status quo, and of the optimal prices under the COMPTN and logit models. We see that the proposed COMPTN demand model recommends lowering the average prices slightly to take advantage of the volume expansion that would follow. This is unsurprising in that the own price elasticities are close to -4. Finally, there is no significant time trend in any of the three sets of price paths we see. Thus, the prescribed price paths do not recommend what may amount to an unrealistic shift in pricing over time. We find qualitatively similar pattern in the optimal price paths of the remaining 10 brands. In the interest of space, these have been put in the appendix.

Profit Simulations

If we assume that our demand specifications (COMPTN and logit respectively), represent the 'true' underlying demand system, or at least capture the true demand system more accurately than does the firm's current approach, then it should be possible to "simulate" the profit that would be realized from the analysis sample and "predict" the profit in the holdout sample, under each demand specification. Using classical prediction under statistical control to simulate profits in the holdout period would only produce hypothetical profit numbers which aren't observed or realized in the data. Hence, we look to Zhang and Krishnamurthi 2004 (henceforth, ZK2004) who devise a test based on observed data in the holdout sample alone to validate their optimal brand promotion strategies. Their basic idea is simple - identify brands in the holdout sample that through random chance closely follow the optimal strategy and compare their profits against that of all other brands. If the optimal strategy indeed enhances profits, then the same should be borne out in the holdout sample profit comparison.

ZK2004 take a manufacturer's perspective to customize promotions for target brand and their optimization routine maximizes each individual brand's profit independently assuming as given the promotion strategies of all other brands. We note that such a strategy would not apply to our case which necessarily views different demand functions as inter-dependent (and hence, solved for all brands simultaneously). However on closer inspection, we find that we can adapt ZK2004's proposed validation test for our inter-dependent product demand context while retaining their core idea that brands that happen to follow (or deviate minimally from) the proposed optimal path should show higher holdout sample profitability than the rest of the sample.

We identify time-periods (weeks) wherein the set of brand price vectors in the holdout sample deviates minimally from those in the optimal solution. We then compare profits in the identified weeks against those in the other weeks. To ensure that any week to week variation arising due to other unforeseen factors (e.g. seasonality) do not confound our results, we normalize our profits by volume. Further, we check for any systematic differences in non-Price Marketing mix variables (primarily, promotion and distribution) between the chosen 8 weeks and the remaining 15 weeks that might systematically impact outcomes and found no such differences. The exact steps we followed were (i) Identify $\sim 1/3$ of the holdout sample's time periods (~ 8 weeks) that most closely follow the optimal strategy's price pattern. (ii) Next, compare the observed average profits across these 8 weeks with those for the remaining 15 weeks. Expectedly, we obtain a different set of 8 weeks for the proposed nDP model compared to the benchmark logit model. The below table displays the results of this validation test.

	\$ average profit per week:	
Holdout sub-samples	nDP optimal Path	Logit based Optimal path
8 weeks closest to optimal path	51088	49543
Remaining 15 weeks	32942	33767

The results say that: (i) following the proposed nDP based linear demand model's optimal price path yields on average a higher profit (~\$51,088 a week) compared to the retailer's current pricing strategy in the remaining 15 weeks; (ii) following the benchmark logit demand model's optimal price path yields on average a higher profit (~\$49,543 a week) compared to the retailer's current pricing strategy in the remaining 15 weeks; and (iii) the proposed nDP based demand model does better than the benchmark logit model in the holdout optimal strategy per week profits.

In sum, conditional on our demand models better capturing the demand system than the firm's current approach, we instances of significant outcome gains from adopting the suggested pricing strategy in a category management context. The recommended price path further shows strong patterns of similarity to current pricing in terms of price-tiers thereby ensuring that the price paths are realistic and practical. We acknowledge that we have not modeled manufacturer reactions to retailer price changes in this paper. However, since there is no change in the price-tiers or in the rankings of brands by average price in the optimal price vectors prescribed by the proposed demand model, concerns about manufacturer and consumer backlash are mitigated. Consumers (and manufacturers) are not treated to unexpected or unusual changes in retail price. This ensures that the prescribed price paths remain practical and implementable from a retailer's perspective.

CONCLUSION

In conclusion, we contribute to the marketing literature on category management in substantive ways. First, we address a gap in the extant empirical literature by accounting for both inter-product and inter-temporal demand dependencies. We do so via incorporating the dynamic elements of competition in a robust, flexible, and scalable fashion in demand models that meet practical considerations and bear face-validity. Second, we develop a parsimonious, flexible and scalable competition construct within a linear demand model framework which allows for SKU-specific heterogeneity in inter-product and inter-period competition effects (in contrast to the baseline mixed logit). Third, we find a Hi-Lo pricing strategy turns out to be optimal in a multi-product, multi-period setting in a real world category. And finally, conditional on our COMPTN demand model better capturing true demand than the focal firm's current approach, our results show that optimizing prices by considering the competitive dynamic effects yields a significant rise in profitability than the focal retailer's current EDLP policy. The mixed logit demand model also performs well and yields a higher category profit than the current pricing policy.

Our research has limitations that future research could address. First, we acknowledge that the model currently does not include manufacturer reactions to retailer price changes. An attempt to expand the model by modeling the game involving manufacturer, retailer and consumer incentives and constraints is an interesting extension and it provides a fruitful avenue for future research. Further, we assume consumer reactions in the aggregate to be in line with those predicted by the proposed demand model. Second, although we have optimized SKU prices, one area for future research is to conduct a field experiment to establish the improvement in profitability at a retail chain (Kumar, Petersen, and Leone 2010). Third, we consider data from one category only whereas data from multiple categories would help generalize the results presented in the paper. We chose the beer category for demonstrating the model and approach

primarily because it has the advantage that the retailer does not have a store brand in the category and therefore avoids potential biases (store brands being treated systematically differently by the retailer than other brands) and confounds (the objective function could be market share maximization with store brands rather than profit maximization). Further, we analyzed data from one retailer only. While this helps maintain consistency in the retailer's core clientele and the product assortment, for more generality, using data from more categories or retailers from different geographic regions would be another avenue of future research. Fourth, among the different marketing mix elements, we optimized price alone. In principle, we could extend our approach to optimize other marketing mix variables such as promotions and distribution. Finally, we used a simplification in our approach in that while our competition function readily computed SKU level sales impact of every change in any SKU's prices, we constrained all SKUs belonging to a brand to have a common, optimal normalized unit price. A more comprehensive model might relax this constraint. Finally, future research could extend our results using a stochastic dynamic programming approach where the error term is incorporated as a state variable in the optimization phase. Thus, although work remains to be done, the concept of using a dynamic, parsimonious description of competitive effects in a category management setting seems viable and worthy of the effort required to understand it more fully.

REFERENCES

- ACNielsen (2004), 13th Annual survey of trade promotion practices. Chicago: ACNielsen.
- Ailawadi, Kusum, Praveen K. Kopalle, and Scott A. Neslin (2005), "Predicting Competitive Response to a Major Policy Change: Combining Game Theoretic and Empirical Analyses," *Marketing Science*, 24 (1), 12-24.
- Antoniak, C.E. (1974), "Mixtures of Dirichlet Processes with Applications to Nonparametric Problems", *Annals of Statistics*, 2(4), pp.1152—1174.
- Basuroy, S., Murali K. Mantrala, and Rockney G. Walters (2001), "The impact of category management on retail price and performance: Theory and evidence", *Journal of Marketing*, 65, 16–32 (October).
- Ben-Akiva M., Steven R. Lerman (1985), "Discrete choice analysis", The MIT Press, Cambridge Massachusetts.
- Berry, S., James Levinsohn, Ariel Pakes (1995), "Automobile prices in market equilibrium", *Econometrica*, 63(4) 841-890.
- Besanko, D., Dubé, J. P., & Gupta, S. (2005), "Own-brand and cross-brand retail pass-through", *Marketing Science*, 24(1), 123-137.
- Carpenter, G.S. and Kent Nakamoto (1989), "Consumer Preference Formation and Pioneering Advantage", *Journal of Marketing Research*, 26(3), pp. 285-298.
- Chen, Y., Hess, J. D., Wilcox, R. T., & Zhang, Z. J. (1999), "Accounting profits versus marketing profits: A relevant metric for category management", *Marketing Science*, 18(3), 208-229.
- Chintagunta, P. K. (2002), "Investigating category pricing behavior at a retail chain", *Journal of Marketing Research*, 39(2), 141-154.
- Conley, T.G., Christian B. Hansen, Robert E. McCulloch, and Peter E. Rossi. (2008), "A Semi-parametric Bayesian Approach to the Instrumental Variable Problem", *Journal of Econometrics*, 144(1), 276-305.
- Dhar, S. K., Hoch, S. J., & Kumar, N. (2001). Effective category management depends on the role of the category. *Journal of Retailing*, 77(2), 165-184.
- Du, R., Lee, E., & Staelin, R. (2005), "Bridge, focus, attack, or stimulate: Retail category management strategies with a store brand", *Quantitative Marketing and Economics*, 3(4), 393-418.

- Eliashberg, J. and Rabikar Chatterjee (1985), "Analytical Models of Competition with Implications for Marketing: Issues, Findings, and Outlook", *Journal of Marketing Research*, 22(3), pp. 237-261.
- Erdem, T. (1996). A dynamic analysis of market structure based on panel data. *Marketing Science*, 15(4), 359-378.
- Fader, P. S., & Hardie, B. G. (1996). Modeling consumer choice among SKUs. *Journal of Marketing Research*, 442-452.
- Fader, P. S., & Lodish, L. M. (1990). A cross-category analysis of category structure and promotional activity for grocery products. *Journal of Marketing*, 52-65.
- Gajanan, S., Suman Basuroy, and Srinath Beldona (2007), "Category management, product assortment, and consumer welfare", *Marketing Letters*, 18, 135–148.
- Gangwar M, Kumar N and Rao RC (2014), Consumer Stockpiling and Competitive Promotional Strategies. *Marketing Science*, 33(1):94-113.
- Gooner, R.A., Neil A. Morgan, and William D. Perreault Jr (2011), "Is Retail Category Management Worth the Effort (and Does a Category Captain Help or Hinder)?", *Journal of Marketing*, 75, 18-33.
- Gruen, T. W., & Shah, R. H. (2000), "Determinants and outcomes of plan objectivity and implementation in category management relationships", *Journal of Retailing*, 76(4), 483–510.
- Guadagni, P.M. and John D. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data" *Marketing Science*, 2(3), 203-238.
- Hall, J. M., Kopalle, P. K., & Krishna, A. (2010), "Retailer dynamic pricing and ordering decisions: category management versus brand-by-brand approaches", *Journal of Retailing*, 86(2), 172-183.
- Hoch, S.J, Xavier Dreze, and Mary E. Purk (1994), "EDLP, Hi-Lo, and Margin Arithmetic", *Journal of Marketing* , 58(4), 16-27.
- Hoch, S., Byung-Do Kim, Alan L. Montgomery, and Peter E. Rossi (1995), "Determinants of store-level price elasticity", *Journal of Marketing Research*, 32, 17–29.
- Kannan, P.K. and Chi Kin Bennett Yim (2001), "An investigation of the impact of promotions on across-submarket competition", *Journal of Business Research*, 53(3), 137-149.
- Kannan, P.K., Barbara Kline Pope, and Sanjay Jain (2009), "Pricing Digital Content Product Lines: A Model and Application for the National Academies Press", *Marketing Science*, 28(4), 620-636.

- Kannan, P. K., & Wright, G. P. (1991). Modeling and testing structured markets: A nested logit approach. *Marketing Science*, 10(1), 58-82.
- Khan, R., Lewis, M., & Singh, V. (2009). Dynamic customer management and the value of one-to-one marketing. *Marketing Science*, 28(6), 1063-1079.
- Kopalle P.K., Ambar G. Rao, and Joao L. Assuncao (1996), "Asymmetric Reference Price Effects and Dynamic Pricing Policies", *Marketing Science*, 15(1), 60-85.
- Kopalle, Praveen, P.K. Kannan, Lin Bao Boldt and Neeraj Arora (2012), "The Impact of Household Level Heterogeneity in Reference Price Effects on Optimal Retailer Pricing Policies," *Journal of Retailing*, 88 (1), 102–14.
- Kumar, V., Fan, J., Gulati, R., & Venkat, P. (2009). Practice Prize Paper-Marketing-Mix Recommendations to Manage Value Growth at P&G Asia-Pacific. *Marketing Science*, 28(4), 645-655.
- Kumar.V. and Robert P. Leone (1988), "Measuring the Effect of Retail Store Promotions on Brand and Store Substitution", *Journal of Marketing Research*, 25(2), 178-185.
- Kumar, V., J. Andrew Petersen, and Robert P. Leone (2010), "Driving Profitability by Encouraging Customer Referrals: Who, When, and How", *Journal of Marketing*, 74 (September), 1-17.
- Levy, Michael, Dhruv Grewal, and Praveen K. Kopalle (2004), "Emerging Trends in Retail Pricing Practice: Implications for Research," *Journal of Retailing*, 80 (3) xiii-xxi.
- Lilien, Gary L., Philip Kotler, and K. Sridhar Moorthy. (1992) *Marketing models*. Englewood Cliffs, NJ: Prentice-Hall
- Ma, Yu, Kusum Ailawadi, Dinesh Gauri, and Dhruv Grewal (2011), "An Empirical Investigation of the Impact of Gasoline Prices on Grocery Shopping Behavior," *Journal of Marketing*, 75 (2), 18-35.
- Mazumdar, T., S.P. Raj, and Indrajit Sinha. (2005), "Reference Price Research: Review and Propositions", *Journal of Marketing*, 69(4), 84-102.
- McAlister, Leigh (2005), "Cross-brand pass-through? Not in Grocery retailing", Marketing Science Institute Report #05-0113.
- McAlister, L. (2007), "Comment-Cross-Brand Pass-Through: Fact or Artifact?,". *Marketing Science*, 26(6), 876-898.
- Moorthy, S. (2005), "A general theory of pass-through in channels with category management and retail competition", *Marketing Science*, 24(1), 110-122.
- Narasimhan C (1988) Competitive promotional strategies. *Journal of Business* 61(4):427–449.

- Newton, M. A., and A.E. Raftery (1994), "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap", *Journal of the Royal Statistical Society, Series B.* 56, 3-48.
- Raju, Jagmohan S. (1992). "The effect of price promotions on variability in product category sales," *Marketing Science*, 11, 2, 207–220.
- Raju, J. S., Sethuraman, R., & Dhar, S. K. (1995), "The introduction and performance of store brands", *Management Science*, 41(6), 957-978.
- Raju JS, Srinivasan V, Lal R (1990) The effects of brand loyalty on competitive price promotional strategies. *Management Sci.* 36(3):276–304.
- Rao RC (1991) Pricing and promotions in asymmetric duopolies. *Marketing Sci.* 10(2):131–144.
- Rao, V. R., Agarwal, M. K., & Dahlhoff, D. (2004). How is manifest branding strategy related to the intangible value of a corporation?. *Journal of Marketing*, 68(4), 126-141.
- Rodríguez A., David B. Dunson, and Alan E. Gelfand. (2008), "The nested Dirichlet process", *Journal of the American Statistical Association*, 103(483), 1131-1154.
- Russell, G.J. and Wagner A. Kamakura (1994), "Understanding brand competition using micro and macro scanner data", *Journal of Marketing Research*, 31(2), 289-303.
- Rust, J. (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher", *Econometrica*, 55(5), 999-1033.
- Seetharaman, P. B., Ainslie, A., & Chintagunta, P. K.. (1999) "Investigating Household State Dependence Effects across Categories". *Journal of Marketing Research*, 36(4), 488–500.
- Shah, Denish, V. Kumar and Yi Zhao (2014), "Diagnosing brand performance: Accounting for the dynamic impact of product availability with aggregate data", *Journal of Marketing Research*, 52(2), pp. 147-165..
- Sriram, S., Subramanian Balachander, and Manohar Kalwani (2007), "Monitoring the Dynamics of Brand Equity Using Store-Level Data, *Journal of Marketing*, 71 (April) 61-78.
- Sudhir, K. (2001), "Competitive Pricing Behavior in the Auto Market: A Structural Analysis", *Marketing Science*, 20(1), pp. 42-60.
- Sun, B., Scott A. Neslin, and Kannan Srinivasan (2003), "Measuring the Impact of Promotions on Brand Switching When Consumers Are Forward Looking," *Journal of Marketing Research*, 40 (November), 389–405.
- Tsiros, M. and David M. Hardesty (2010), "Ending a Price Promotion: Retracting It in One Step or Phasing It Out Gradually", *Journal of Marketing*, 74(1), 49-64.

Wedel, M. and Jie Zhang (2004), "Analyzing Brand Competition Across Subcategories", *Journal of Marketing Research*, 41(4), 448-456.

Varian HR (1980) A model of sales. *American Economic Review* 70(4): 651–659.

Voleti, S. and Pulak Ghosh (2013), "A robust approach to measure latent, time-varying equity in hierarchical branding structures", *Quantitative Marketing and Economics*, 11 (3), 289-319.

Voleti, S., Praveen K. Kopalle and Pulak Ghosh (2015), "An Inter-Product Competition Model Incorporating Branding Hierarchy and Product Similarities Using Store Level Data", *Management Science* (Forthcoming).

Zenor, M. J. (1994). The profit benefits of category management. *Journal of Marketing Research*, 202-213.

Zhang, J., & Krishnamurthi, L. (2004). Customizing promotions in online stores. *Marketing Science*, 23(4), 561-578.

Figure 1: Flowchart of the Proposed Approach (Numbered by sequence of implementation)

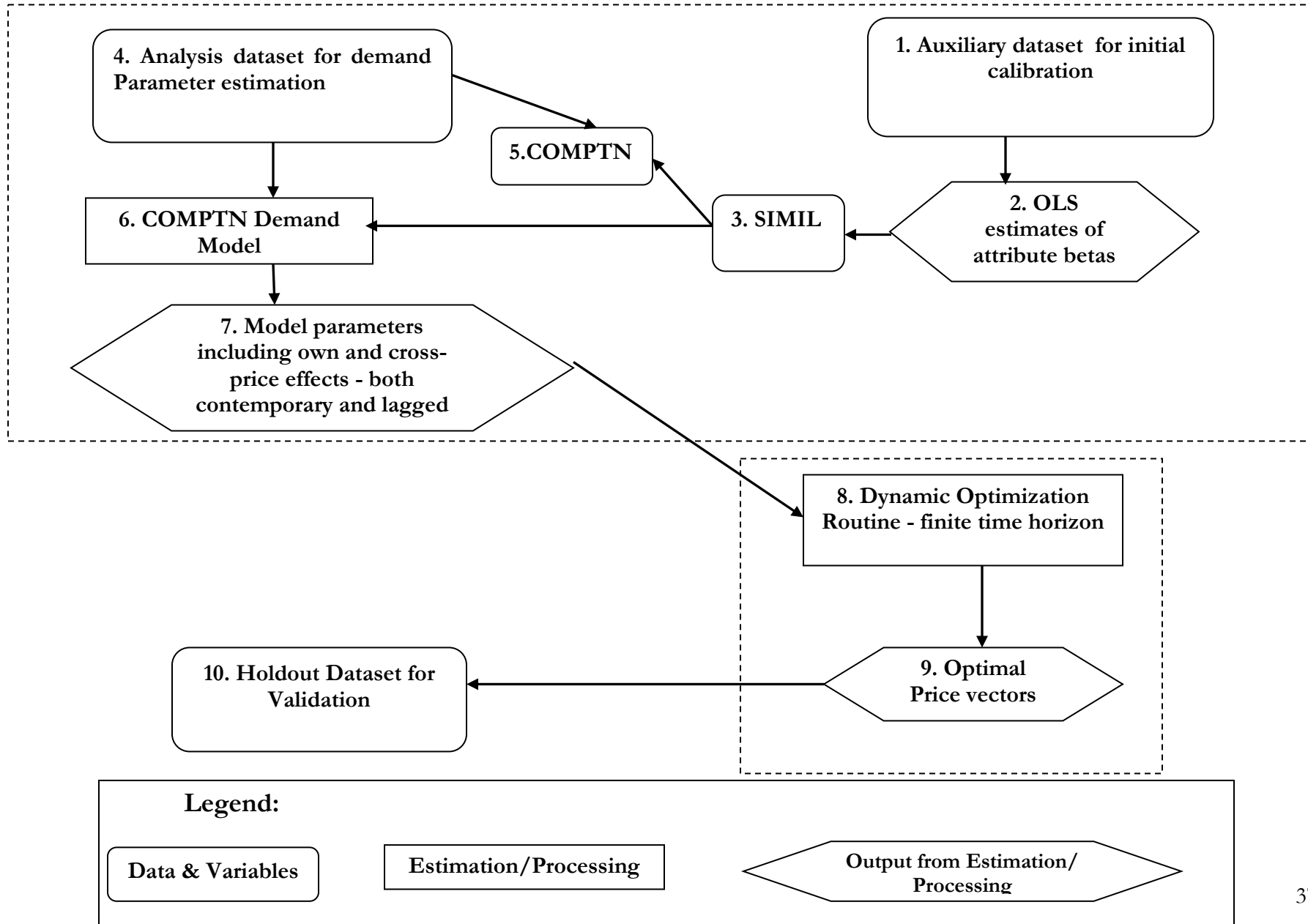


Table 1. Positioning the paper amid Related Literature

Studies focussed on Profit Maximization under Category Management	Accounts for inter-product dependence at the SKU level	Accounts for inter-temporal demand dependence	Yields optimal category profits	Applies to stable assortments in established categories
Zenor (1994)	✗	✗	✓	✓
Basuroy, Mantrala and Walters (2001)	✓	✗	✗	✓
Chintagunta (2002)	✗	✗	✓	✓
Kumar, Fan, Gulati and Venkat (2009)	✗	✗	✓	✓
Mantrala, Seetharaman, Kaul, Gopalakrishna and Stam (2006)	✓	✗	✓	✗
Hall, Kopalle and Krishna (2010)	✗	✓	✓	✓
Shah, Kumar and Zhao (2015)	✓	✓	✗	✗
Voleti, Kopalle and Ghosh (2015)	✓	✗	✗	✓
This paper	✓	✓	✓	✓

Table 2. Summary of Analysis Variables - Beer Category Data from one Retailer

Group	Variable	Description	Mean(s.d.)	Group	Variable	Proportion
Dependent Variable	Sales	Beer Volume in Fl.Oz	39204(61531)	Product Attributes		
	Price	Retail price per fluid ounce	.07(.02)	Beer Type	Ale	.011
Distribution	COST	Wholesale price per ounce	.06(.01)		Lite	.386
	DISTBN	% stores that sold the SKU	.70(.29)		Craft	.095
Promotion	UNITSP	% Units sold on promotion	.22(.36)		Regular	.519
	ADSPEND	National ad-spend (\$mn p.a.)	13.97(20.38)	Beer Color	Light	.679
Control Variables					Amber	.094
MONTH	Jan	If week is in Jan	.170		Dark	.042
	Feb	If week is in Feb	.173	Packaging	Golden	.185
	March	If week is in March	.173		Bottle	.711
	April	If week is in April	.175		6P 12Oz	.299
	May	If week is in May	.175		12P 12Oz	.454
	June	If week is in June	.134		18 Pack	.013
					24 Pack	.021

Table 3: Asymmetric Marginal Effects Implied by the COMPTN measure for 3 SKUs

		SKU i (Rival SKU)		
		SKU #56	SKU #49	SKU #42
SKU j (Focal SKU)	SKU #56	-4.112 [SIMIL (j, i)] { COMPTN(j, i) }	2.4014 [1.8660] { 10.4977 }	0 [0] { 0 }
	SKU #49	3.052 [1.8660] { 9.711 }	-4.026 [SIMIL (j, i)] { COMPTN(j, i) }	0.6791 [.0812] { .42260 }
	SKU #42	0 [0] { 0 }	0.4036 [.0812] { .3492 }	-4.572 [SIMIL (j, i)] { COMPTN(j, i) }

Table 4: Comparative Fit and Prediction due to SIMIL & COMPTN

Model Specification	Calibration Sample (Model fit)		Holdout Sample (Predictive fit)	
	RMSE	Log Marginal Density	RMSE	Log Marginal Density
Both SIMIL and COMPTN	0.1661	566.378	0.217	-81.48
SIMIL only	0.1804	533.405	0.2371	-113.121
COMPTN only	0.1792	559.906	0.2331	-60.673
Neither SIMIL nor COMPTN	0.184	518.696	0.2386	-110.127

Figure 2: Variance Explained attributable to Groups of Regressors

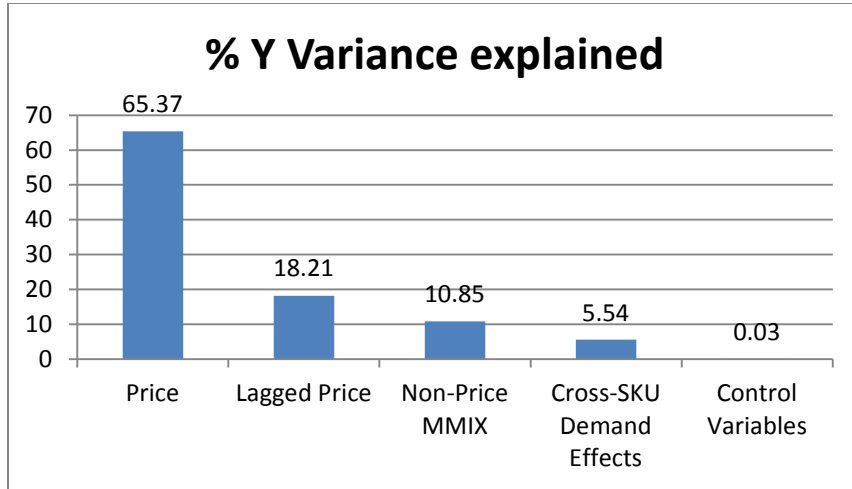


Table 5: Demand Model Estimates for the Log-Linear COMPTN Specification

Panel A		
Variables	Metrics	Estimates
SKU specific Random Effect		-5.621 [-6.34, -5.08]
Price Parameter		-3.936 [-5.01, -2.87]
Distribution Parameter	Average of Posterior Means [Range of Posterior Means]	.61 [-.10, 1.23]
Promotion Parameter		.083 [-.09, .15]
Competition Response		-1.286 [-1.93, -.68]
Panel B		
Variables	Metrics	Estimates
Ln ADSPEND		.037 [.027, .045]
SIMIL		2.192 [1.919, 2.451]
Lagged Ln PRICE	Posterior Mean [95% Credibility Interval]	-2.054 [-2.154, -1.958]
Lagged Ln DISTBN		.007 [-.041, .055]
Lagged Ln PROMO		-.005 [-.011, .002]

Table 6: Parameter Estimates for the Baseline Mixed Logit Model

Contemporaneous MMIX Variables	Metric	Estimates
SKU Specific		-5.43
Random Effect		[-9.17, -.29]
Price Parameter	Average of Posterior Means [Range of the Posterior Means]	-34.82 [-79.36, -16.04]
Distribution Parameter		1.065 [-.06, 2.96]
Promotion Parameter		.016 [-.11, .09]
1 period Lagged MMIX Variables		
Lagged Price Parameter		-.7184 [-1.23, .13]
Lagged Distribution Parameter	Posterior Mean [95% Credibility Interval]	.182 [.09, .43]
Lagged Promotion Parameter		.0037 [.0016, .0068]
Model Fit and Inference		
Calculated Nonparametric	In-sample Ln Sales RMSE	.159
Measure of fit	Holdout-sample Ln Sales RMSE	.223

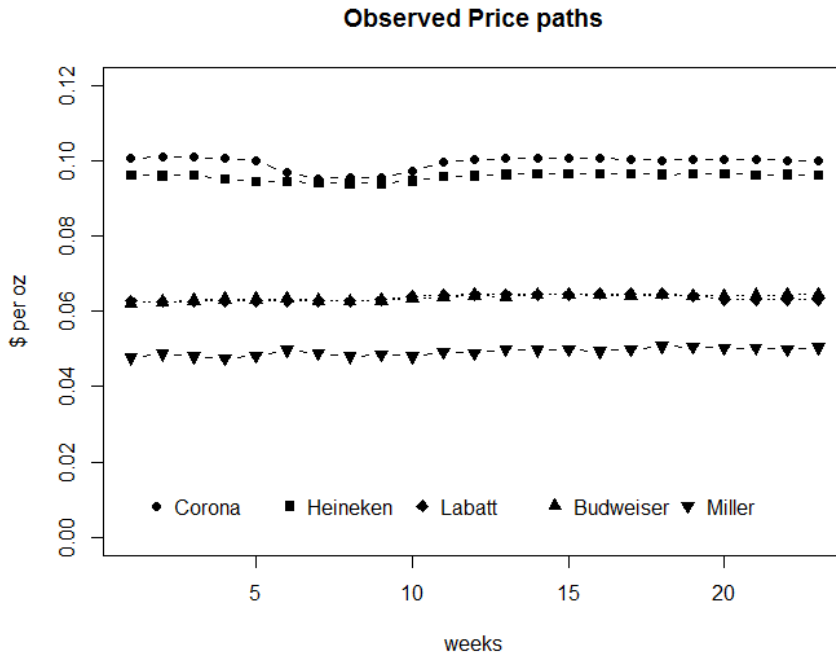
Table 7: Correlation between Different Price paths

	Current Price Path	Optimal Path from COMPTN Model	Optimal Path from Logit Model	Price Summary (\$/oz) Mean (s.d.)
Current Price Path	1	.26*	-.05*	.076 (.0149)
Optimal Path from COMPTN Model	.3	1	-.1*	.074 (.018)
Optimal Path from Logit Model	.38	0.47	1	.076 (.016)

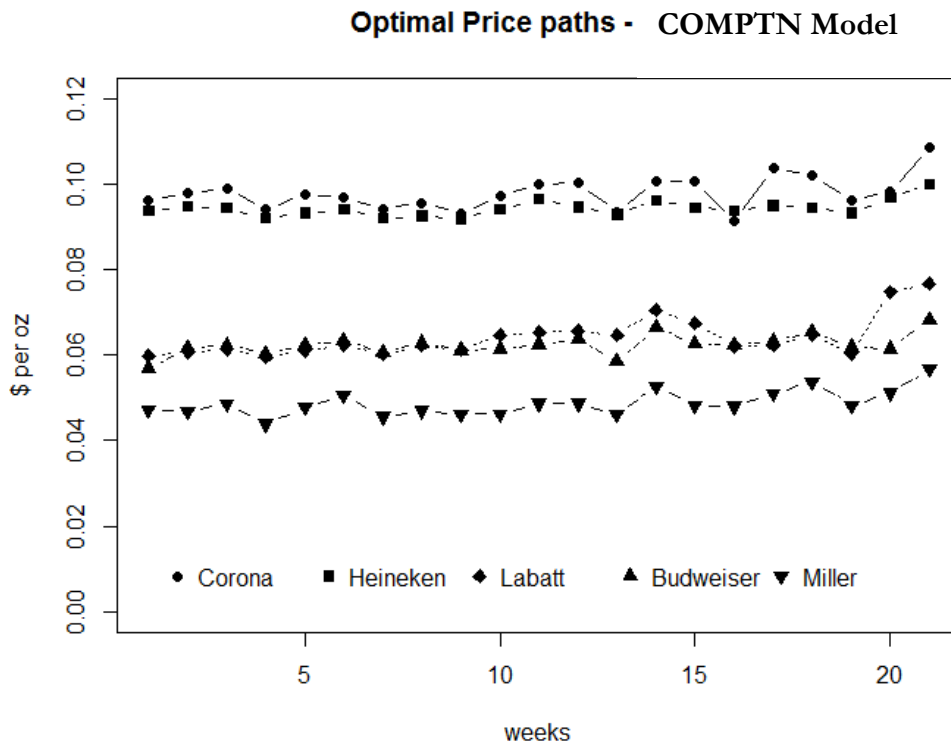
* Spearman correlation coefficients; Lower triangular shows Pearson Correlation coefficients.

Figure 3: Current and Optimal Price Paths by Brand (\$Price per ounce Vs Week)

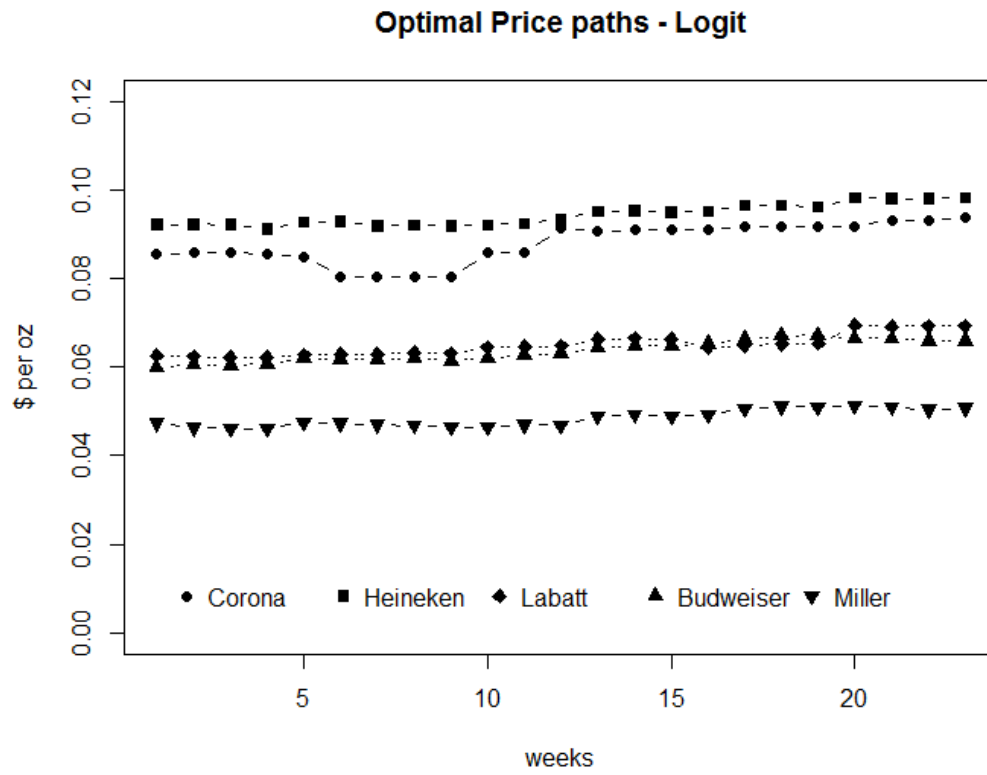
(A) Current Pricing Path



(B) Optimal Pricing Path under the COMPTN Model

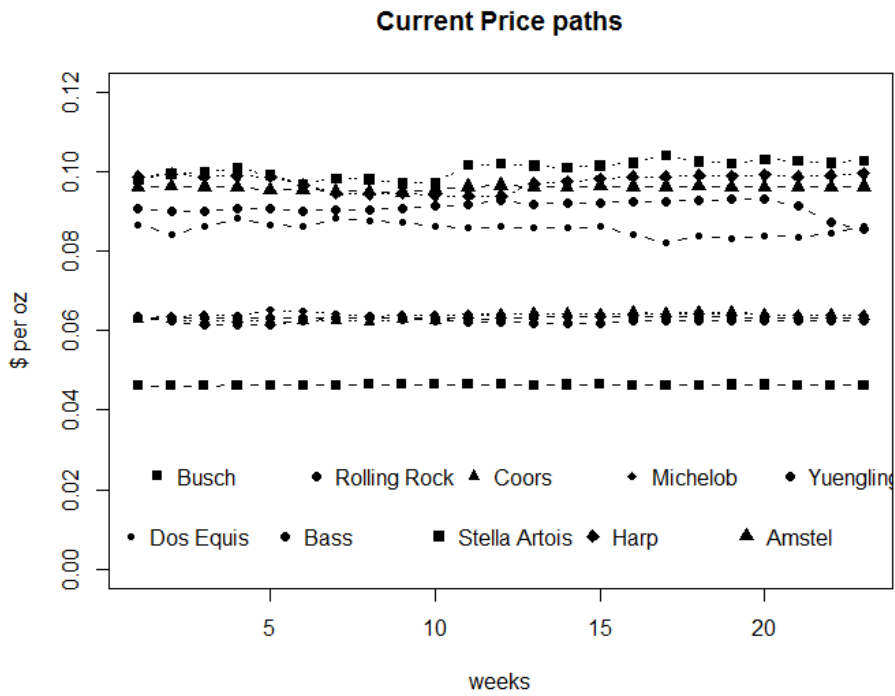


(C) Optimal Pricing Paths under the Mixed logit Model

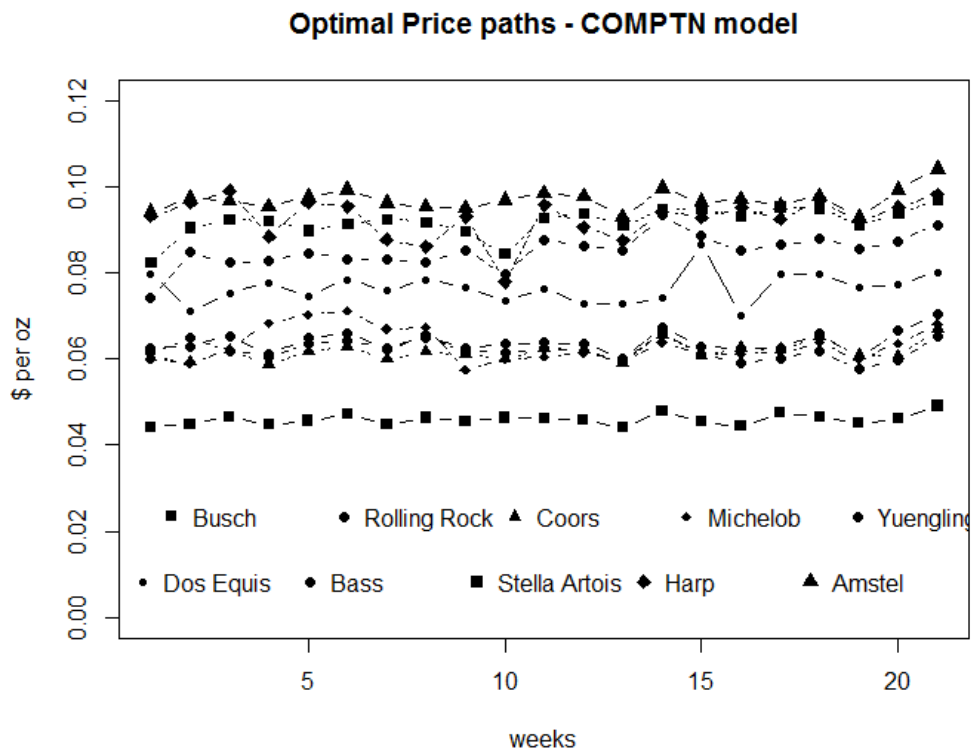


APPENDIX: Figure A1 - Price Paths for the Remaining 10 brands

(A) Current Pricing Paths



(B) Optimal Pricing Path under the COMPTN Model



(C) Optimal Pricing Paths under the Mixed logit Model

