Marketing Science Institute Special Report 09-213

The Long-term Effect of Marketing Strategy on Brand Sales

Berk Ataman, Harald van Heerde, and Carl Mela

The Long-term Effect of Marketing Strategy on Brand Sales

M. Berk Ataman
Assistant Professor of Marketing
Rotterdam School of Management, Erasmus University
P.O. Box 1738, 3000 DR, Rotterdam, Netherlands
Phone: +31-10-408-2612
Fax: +31-10-408-9011
bataman@rsm.nl

Harald J. van Heerde
Professor of Marketing
Waikato Management School, University of Waikato
Private Bag 3105, Hamilton 3240, New Zealand
Phone: +64-7-838-4089
heerde@waikato.ac.nz

Carl F. Mela*
Professor of Marketing
Fuqua School of Business, Duke University
1 Towerview Drive, Durham, NC 27708
Phone: +1-919-660-7767
Fax: +1-919-660-7790
mela@duke.edu

* The authors would like to thank IRI and TNS Media Intelligence for providing the data, Marketing Science Institute and Zyman Institute for Brand Science for research support. Ataman and Van Heerde would like to thank Netherlands Organization for Scientific Research for research support. Earlier versions of this paper benefited from valuable comments of seminar participants at Northwestern University, Yale School of Management, Erasmus University Rotterdam, University of Groningen, Catholic University Leuven, Free University Amsterdam, and Tilburg University.

The Long-term Effect of Marketing Strategy on Brand Sales

Report Summary

Few studies consider the relative role of the entire marketing mix on long-term performance of mature brands –instead emphasizing advertising and price promotion. Hence, little guidance is available to firms regarding the relative efficacy of their various marketing expenditures over the long run. Accordingly, the authors supplement prior research by considering the long-term effect of the entire marketing mix (advertising, price promotion, product, and place) over a large number of categories. To do this, the authors combine five years of advertising and weekly chain-level scanner data for 25 product categories and 70 brands in the four largest retail chains in France. Using a multivariate dynamic linear transfer function model, the authors find that the total (short-term plus long-term) sales elasticity for product is 1.37 and for distribution it is .74. In sharp contrast, the total elasticities for advertising and discounting are only .13 and .04, respectively. This result stands in marked contrast to the previous emphasis in the literature on price promotions and advertising.

Keywords: Advertising, Price promotion, Distribution, Product, Long-term effects, Brand Performance, Bayesian Time Series Methods, Dynamic Linear Models, Empirical Generalizations

Firms annually spend hundreds of billions of dollars to implement their marketing strategy. Much headway has been made explaining how these expenditures enhance brand performance over the short term (Bucklin and Gupta 1999)[1]. More recently, attention has been focused on the longer-term effect of marketing strategy on brand performance, particularly with respect to price and promotion (e.g., Boulding, Lee, and Staelin 1994; Jedidi, Mela, and Gupta 1999; Nijs et al. 2001; Pauwels, Hanssens, and Siddarth 2002; Steenkamp et al. 2005). Yet, there has been little emphasis on the effects of product (e.g., line length) and place (e.g., distribution breadth) on brand performance. Accordingly, a critical question remains unanswered (Aaker 1991; Ailawadi, Lehman, and Neslin 2003; Yoo, Donthu, and Lee 2000): which elements of the marketing mix are most critical in making brands successful?

[FIGURE 1 AND FIGURE 2 ABOUT HERE]

To illustrate these points, we show in Figure 1 and Figure 2 the historical performance of two brands over a five-year period; one that contracted dramatically (Brand C, C = Contracted), and one that grew considerably (Brand G, G = Grew). Figure 1 and Figure 2, respectively, show sales volume, promotion activity, advertising spending, distribution breadth, and product line length for Brand C and Brand G over time. The brands and variables are from a data set that we discuss in more detail in subsequent sections. Comparison of sales volume between the first and second half of the data reveals a considerable 60% sales contraction for Brand C, which contrasts to an 87% growth for Brand G. This difference in performance begs the question of what strategies discriminate between the performances of these brands.

To attain insights into this question, we first consider Brand C. One possibility is that it is experiencing a mature brand's cycle-like fortunes (Pauwels and Hanssens 2007). Yet, closer inspection reveals a link between its performance and its marketing strategy. Brand C's downward sloping sales (Figure 1a) during its first four years coincide with frequent and deep discounting (Figure 1b), negligible advertising (Figure 1c), lower distribution (Figure 1d), and shorter product line (Figure 1e). Of note, its

[1] By short-term, we mean the immediate effect of marketing on current week's sales. In contrast, long-term refers to the effect of repeated exposures to marketing over quarters or years.

sales turn around in the last year of our data. This period is characterized by increased product variety, distribution, and advertising, while discounting was curtailed.

Brand G's sales (Figure 2a) show a marked increase shortly after week 100. This might illustrate the (autonomous) take-off of a small brand (Golder and Tellis 1997). Once again, a link between brand performance and its marketing strategy can be established. The increase in sales coincides with heavy product activity (Figure 2e), high advertising spending (Figure 2c), increased distribution (Figure 2d), and diminished price promotions (Figure 2b).

Together, these examples suggest that product, distribution, and advertising seem to enhance brand performance, while discounts do little in the way of brand building. Yet, these cases are anecdotal (and involve only two categories) and the various mix effects are confounded. In fact, the correlation between these strategies suggests that it is especially important to consider them in unison otherwise an assessment of effects in isolation might lead one to attribute a brand's success to the wrong strategy. For example, this would occur if advertising was correlated with distribution. In this event, increases in sales might mistakenly be attributed to advertising in the absence of distribution. Accordingly, our objective is to investigate more systematically how marketing affects brand performance in the long run. By analyzing the weekly performance of brands in 25 categories over five years, we identify the marketing mix strategies that correlate most highly with growth in brand sales and more potential to command higher prices.

Our results substantiate the belief that distribution and product decisions play a major role in the long-term performance of brands. By computing the relative long-term sales elasticities of the various marketing strategies we find that product effects are 63% and distribution effects are 30%. In contrast, the effect of advertising and discounting are only 6% and 1%, respectively. Moreover, while the long-term effect of discounting (negative) is about one-third of the magnitude of its short-term effect (positive), the long-term effects of the other marketing variables tend to be 4 to 16 times their short term effects, testifying to the long-term role they play in brand performance. Also of note, the total (long-term plus short-term) elasticities of line length and distribution breadth are more substantial (1.37 and .74

respectively) than the advertising and discount elasticities (.13 and .04 respectively). These results illustrate discounts do little to build a brand over the long term.

These findings arise from the application of a multivariate Dynamic Linear Model (DLM) that links brand sales to marketing strategy. The approach offers a flexible means for assessing how marketing affects intercepts and sales response parameters (e.g., elasticities) over time. Moreover, the approach (i) controls for endogeneity in pricing and marketing variables, (ii) partials the role of past performance from marketing spend, and (iii) considers competitive interactions in marketing. To our knowledge, the DLM has never been applied to a problem of this scale, making this application a notable advance.

The paper is organized as follows. We first discuss the literature on long-term effects of the marketing mix on brand performance. Second, we discuss theories pertaining to how the marketing mix affects brand performance in the long run. Third, we develop our model and overview estimation. Fourth, we describe the data and variables. Fifth, we present the results. Last, we conclude with a summary of findings and future research opportunities.

LITERATURE ON LONG-TERM EFFECTS OF THE MARKETING MIX

Table 1 samples the current state of the long-term effects literature and indicates (i) a prevalent focus on certain marketing instruments, (ii) the existence of various brand performance measures, and (iii) a clear divide between modeling approaches. We address these issues subsequently and highlight our points of difference and parity.

[Table 1 ABOUT HERE]

First, Table 1 indicates that most studies focus on promotion and advertising instead of distribution and product. Hence, these studies a) cannot provide insights into the relative effects of marketing variables and b) risk suffering omitted variable bias as these strategies can be correlated. Related, our personal interviews with senior research managers at different consumer packaged goods firms yielded a similar focus regarding the prevalence of advertising and discounting in industry research. Yet, these managers remain unclear regarding whether this attention is misplaced in the sense that product and distribution actually do play a greater role in brand performance. Another reason for the focus on advertising and promotion in industry might pertain to their ease of measurement: it is easy to observe the immediate effect of deals on sales, but much harder to assess how product innovation affects brands in the long-term. Immediacy may also play a role, as the short-term effect of a discount is large while the effect of building distribution may take some time. As brand managers are promoted quickly, there is little incentive to invest in long-term brand building. This underscores the importance of tools to measure the longer-term effects of marketing strategy on brands, lest the emphasis on short-term metrics induces brands to weaken over time. It further highlights the relevance of a large scale systematic study to determine whether the industry focus on discounting might be misplaced. Accordingly, the question "How does the marketing mix influence brand equity in the long run?" has been a top research priority of the Marketing Science Institute ever since 1988 (MSI research priorities 1988-2008). One reason that this question has been around for so long is that answering it requires the combination of very extensive data sets and a methodology that is able to measure long-term effects while coping with the common challenges of empirical modeling such as (i) endogeneity in marketing, (ii) performance feedback (e.g., the effect of past sales on current marketing expenditures), and (iii) competitive interactions. This research meets these challenges as discussed in the following sections.

A second observation from Table 1 is that these studies differ in their use of brand performance measures. Brand performance or brand equity has been conceptualized and operationalized using stock market returns (Simon and Sullivan 1993), brand attitudes (Aaker 1991), and brand sales or choice data (Ailawadi, Lehmann and Neslin 2003). Though each has its respective benefits, most studies in Table 1 fit in the third stream as does ours.

Papers embedded in this stream commonly propose different measures for brand equity. The first measure suggests assessment of brand equity through base sales, which is operationalized as the brand intercept in a sales model (Kamakura and Russell 1993; Kopalle, Mela, and Marsh 1999).[2] We adopt this operationalization and contend that a brand performs better as its base sales grow. The second measure

---

[2] Base sales are a brand's sales when all marketing variables are at their means. This is different from baseline sales, which is sales in the absence of a promotion.

pertains to the notion that well-differentiated brands can command higher regular prices and margins than otherwise similar goods (Swait et al. 1993). As price premiums are inversely related to the brand's regular price elasticity (Boulding et al. 1994; Nicholson 1972)[3], we adopt regular price elasticity as our second measure of brand performance. Consistent with the brand differentiation view, we consider a brand strong when the regular price elasticity is high (i.e., close to zero). Importantly, the effect of marketing strategy may differ materially between base sales and price elasticity. For example, distribution may enhance the base sales of a good, but enhanced availability might encourage "cherry picking" across stores leading to enhanced price response (Fox and Hoch 2005). We elaborate on these differences in the next section.

Table 1 also reveals that studies that do consider product and distribution in addition to advertising and discounting (e.g., Ataman, Mela, and Van Heerde 2008) emphasize their effects on sales and not their implications for elasticities. Another point which our study differs from the Ataman et al. (2008) study is that the latter considers new brands only. These brands are qualitatively different from the mature brands we study; mature brands have an installed base of customers and an existing distribution network, which are lacking for new brands.

A third observation from Table 1 is that there are two dominant approaches in modeling the long-term effects of the mix: varying parameter models and VAR models. Whereas inertia in marketing spend (Pauwels 2004) and performance feedback (Horvath et al. 2005) are integral parts of VAR models, they often ignore varying parameter effects. Varying parameters are relevant since marketing strategy may affect both intercepts and price elasticities. In contrast, varying parameter models (including the Bayesian variant "DLM") often ignore inertia and feedback effects yet these are important to calculate the returns accruing from marketing investments over the long term. Therefore, in our application, we combine the two approaches and develop a varying parameter model (DLM variant) for a system of equations that

[3] Assume that brand i faces a multiplicative demand curve, $q_i = \alpha_i p_i^{\beta_{ii}} \prod p_j^{\beta_{ij}}$ for all $j \neq i$, where q_i is demand, p_i is the regular price of brand i, α_i is the intercept, and β_{ij} the price elasticity of brand i to brand j. If the marginal cost of production is c_i, the profit function is given by, $\pi_i = q_i(p_i - c_i)$. Then solving $max\ \pi_i = 0$ for p_i gives profit maximizing price, $p^*_i = c_i/((1/\beta_{ii})+1)$. Hence price, as well as the percent profit margin = $(p_i - c_i)/p_i \equiv -1/\beta_{ii}$, increase as regular price elasticity decreases. Note further that when the demand function is multiplicative competitor's price drops from the first order condition.

considers the role of inertia in marketing spend and performance feedback. Our analysis indicates both inertia and feedback to be quite substantial.

In sum, our study extends the current literature on the long-term effects of marketing strategy on brand performance by (i) considering the full marketing mix, (ii) adopting base sales and price elasticity as performance measures, and (iii) specifying a system of equations with time-varying parameters.

THE EFFECT OF THE MIX ON BRAND PERFORMANCE

The following sections overview the current literature on the long-term effects of price promotions, advertising, distribution, and product on brands, and their relation to base sales and regular price elasticity (see Table 2). We note that our discussion of distribution and product is more tentative given the dearth of work in the area. We then conclude by discussing the relative efficacy of the various marketing strategies.

Price Promotion

While some studies in the literature suggest a negative long-term impact of price promotions on base sales (Foekens, Leeflang, and Wittink 1999; Jedidi, Mela, and Gupta 1999), others suggest the opposite effect due to the positive effects of state dependence (Keane 1997) and purchase reinforcement (Ailawadi et al. 2007). Others have found only a fleeting negative effect (Pauwels, Hanssens, and Siddarth 2002). Overall, it is not clear whether the positive effect dominates the negative effect on base sales, and a large-scale generalization seems necessary.

In contrast, discounting policies are typically found to decrease price elasticities (make them more negative) by focusing consumers' attention to price-oriented cues (Boulding et al. 1994; Mela, Gupta, and Lehmann 1997; Papatla and Krishnamurthi 1996; Pauwels et al. 2002).

Advertising

Brand-oriented advertising (e.g., non-price advertising) strengthens brand image, causes greater awareness, differentiates products and builds brand equity (Aaker 1991; Keller 1993). Advertising may also signal product quality leading to an increase in brand equity (Kirmani and Wright 1989). Accordingly, several authors have found advertising to have a positive and enduring effect on base sales (e.g., Dekimpe and Hanssens 1999).

Two different schools of thought in economic theory, namely information and market power theories, offer alternative explanations for the impact of advertising on price elasticity. Information theory suggests that advertising may increase competition by providing information to consumers about the available alternatives, thus making price elasticities more negative, whereas market power theory argues that advertising may increase product differentiation, thus making price elasticity less negative (Mitra and Lynch 1995). Related, Kaul and Wittink (1995) indicate that brand-oriented advertising increases price elasticity while price-oriented advertising decreases it. Mela, Gupta, and Lehmann (1997) note that national brand television advertising is predominantly brand-oriented. Accordingly, we expect national television advertising, as observed in our data, to increase price elasticities (making them less negative).

Product

Similar to advertising, product activity (e.g., innovations, changes in form, etc) enhances a brand's perceived quality, increases purchase likelihood and builds equity (Berger, Draganska, and Simonson 2007). However, research regarding the long-term effect of product on brand performance is quite limited compared to research on the long-term effects of promotions and advertising. Therefore our expectations regarding the effects of extending the product line are tentative. We posit that the long-term effect of increased product line length on base sales is incumbent upon the degree to which cannibalization offsets incremental sales garnered by serving more segments. In general, we argue offering more products has a small but positive effect on base sales because we do not expect cannibalization to entirely offset the increased demand. Accordingly, some studies in the literature suggest that product innovation is positively related to brand performance in the long run (Ataman, Mela, and Van Heerde 2008; Sriram, Balachander, and Kalwani 2007; Pauwels 2004). We expect that more differentiated or customized alternatives increase price elasticity (making it less negative) because strongly differentiated items can serve loyal niches.

Distribution

Distribution breadth (the percent of distribution that carries a brand) can affect brand performance, but as with product, theoretical and empirical evidence for these effects are limited. We expect that increases in the breadth of distribution lead to higher base sales as the wider availability facilitates consumers' ability to find the brand (Bronnenberg, Mahajan, and Vanhonacker 2000).

Two competing expectations can be formulated for the effect of distribution breadth on price elasticity. First, broader distribution may increase the chance of within-brand price comparison across stores, commonly called "cherry picking" (Fox and Hoch 2005). This leads to an increased emphasis on price and an attendant decrease in price elasticity. In contrast, broader distribution signals manufacturer commitment to the brand and potentially its success in the marketplace. A similar signaling effect is also observed for advertising (Kirmani and Wright 1989). Given the competing arguments, we treat the effect of distribution breadth on elasticity as an empirical question. Table 2 summarizes the expected effects of marketing on brand performance.

[TABLE 2 ABOUT HERE]

Relative Effects

Of interest is the relative magnitude of these effects. To our knowledge, no research incorporates all of these effects into a single framework over a large number of categories, so any discussion of the relative magnitude of these effects is necessarily speculative. Complicating this task, marketing strategy is affected by performance feedback, competitor response and inertia. For example, a positive effect on base sales can be amplified in the presence of inertia because the positive effect manifests itself not only in the current period, but subsequent periods as well. There is ample reason to believe some aspects of the mix might be more enduring then others -for example, it takes more time to make changes to the product line than to implement a deal.

Our personal communications with firms and colleagues suggest most individuals expect distribution and product innovation to have the greatest overall long-term effects on brand sales. Distribution and product line length are a necessary condition for sales: no distribution or products imply no sales. Some evidence for this exists for new brands. A recent study by Ataman, Mela, and Van Heerde (2008) shows that distribution plays a central role in building new brands. Product innovation is also likely to have considerable effects as it is a core source of differential advantage. In contrast, advertising and pricing are more limited in their ability to differentiate goods. In sum, we expect product and distribution to matter most for brand performance in the long run.

MODELING APPROACH

Overview

We seek to allow the base sales and regular price elasticity to vary over time as a function of marketing strategy. Dynamic Linear Models (Ataman, Mela, and Van Heerde 2008; West and Harrison 1997) are well-suited to this problem. The general multivariate form of our model is:

(1a) $$Y_t = F_t\theta_t + X_t\eta + Z_t\zeta + \upsilon_t$$

(1b) $$\theta_t = G\theta_{t-1} + Z'_{t-1}\gamma + \omega_t$$

where Y_t is a vector in which the log sales of brand j in chain s at time t is stacked across brands and chains. F_t is a regressor matrix consisting of an intercept and log regular price, while X_t is a regressor matrix including a number of control variables, such as feature/display and seasonality, which affect sales. Z_t includes brands' marketing strategies, specifically advertising expenditures, price discounting, distribution breadth, and product line length. We assume $\upsilon_t \sim N(0,V)$, where V is the covariance matrix of error terms in (1a). The *observation equation* (1a) models the short-term effect of marketing activities on sales. Note that this equation yields period-specific estimates (stacked in θ_t) for intercepts (base sales) and regular price elasticities. We allow these to vary over time as described by the *system equation* (1b) in order to measure the long-term effect of marketing strategies on base sales and regular price elasticity. The system evolution matrix G measures the duration of these strategies –comparable to the decay rate of advertising stock. The stochastic term ω_t are assumed to be distributed $N(0,W)$.

Importantly, the DLM methodology accounts for evolution/nonstationarity in the data. Paraphrasing from West and Harrison (1997, p 299-300), DLM approaches model the original time series directly, without data transformations such as differencing. High levels of non-stationarity cannot usually be removed by differencing or other data transformations, but instead they are directly modeled through a DLM representation. For a detailed discussion on the benefits of DLM methodology and its relation to other time series models (e.g., VAR) see van Heerde, Mela, and Manchanda (2004) and West and Harrison (1997).

Next, we elaborate upon this basic specification (1a) and (1b) and detail how we extend it to control for endogeneity in prices and marketing mix and performance feedback.

Model Specification

Observation equation: Short-term effects. To capture the short-term effect of marketing activity on a brand's sales in a given chain, we operationalize equation (1a) as a log-log model similar to Van Heerde, Mela, and Manchanda (2004) and others,

$$\begin{aligned}
\overline{\ln SALES}_{jst} &= \alpha_{jt} + \beta_{jt}\overline{\ln RPR}_{jst} + \mu_j \overline{\ln PI}_{jst} + \phi_j \overline{FND}_{jst} + \sum_{\substack{j'=1 \\ j'\neq j}}^{J} \rho_{1j'j}\overline{\ln CRPR}_{j'st} \\
&\quad + \sum_{\substack{j'=1 \\ j'\neq j}}^{J} \rho_{2j'j}\overline{\ln CPI}_{j'st} + \tau_{0j} TEMP_t + \sum_{i=1}^{I} \tau'_{ij} HDUM_{it} \\
&\quad + \zeta_{1jt}\overline{ADV}_{jt} + \zeta_{2j}\overline{DSC}_{jt} + \zeta_{3j}\overline{DBR}_{jt} + \zeta_{4j}\overline{LLN}_{jt} + \xi_{jst} + \upsilon^{S}_{jst}
\end{aligned} \tag{2}$$

where ln $SALES_{jst}$ represents the log sales of brand j in chain s in week t, ln RPR_{jst} is the log inflation adjusted regular price, and ln PI_{jst} is the log of price index, which is defined as the ratio of actual price to regular price. FND_{jst} indicates whether there was a feature and/or display without a price discount, while ln $CRPR_{j'st}$ and ln $CPI_{j'st}$ are log cross regular prices and log cross price indices, respectively. $TEMP_t$ is the average temperature in week t and $HDUM_{it}$ is a vector of holiday dummies for events such as Christmas and Easter. The four marketing variables included in equation (2) are advertising expenditure (ADV_{jt}), national discount depth (DSC_{jt}), distribution breadth (DBR_{jt}), and product line length (LLN_{jt}).

We standardize all variables (after taking logs, if applicable) within brand-chain to control for unobserved fixed effects and indicate this by the superscripted bar. This standardization also facilitates comparison of effect sizes across the mix and categories (where price is typically expressed in different equivalency units such as liters or grams) and implies that the model uses within-brands variation over time for inferences.

In equation (2), α_{jt} is the brand specific time varying intercept, which can be construed as base sales since all independent variables have been mean-centered. The time varying brand-specific regular price elasticity coefficient β_{jt} is the second central parameter. We also incorporate a number of control variables in the model; μ_j is the promotional price elasticity, ϕ_j is the feature and/or display log multiplier, $\rho_{1j'j}$ and $\rho_{2j'j}$ are cross regular and promotional price elasticities, and τ_{ij}'s ($i = 0,\ldots,I$) capture seasonal variation. ζ_{kj} ($k = 1,\ldots,4$) capture the short-term (contemporaneous) effects of marketing activities on sales. Given the rich literature on advertising dynamics (e.g., Bass et al. 2007; Naik, Mantrala, and Sawyer 1998), we allow the advertising effect to vary over time (ζ_{1jt}) with a random walk evolution $\zeta_{1jt} = \zeta_{1jt\text{-}1} + \omega^{\zeta}_{jt}$. We include ξ_{jst}, a brand-chain specific intercept, to account for potential first order autoregressive errors ($\xi_{jst} = \lambda^{\xi}_{js}\xi_{jst\text{-}1} + \omega^{\xi}_{jst}$). υ^{S}_{jst} is an error term, which is assumed to be distributed normal and independent across time.

System equation: Long-term effects. A core contention of our research is that brands' base sales and regular price elasticities vary over time as a function of marketing variables. To test these conjectures, we specify the long-term effect of marketing strategies on these two performance measures by operationalizing equation (1b) as follows:

$$\alpha_{jt} = \delta^{\alpha}_{j} + \lambda^{\alpha}_{j}\alpha_{jt-1} + \gamma^{\alpha}_{1}\overline{ADV}_{jt-1} + \gamma^{\alpha}_{2}\overline{DSC}_{jt-1} + \gamma^{\alpha}_{3}\overline{DBR}_{jt-1} + \gamma^{\alpha}_{4}\overline{LLN}_{jt-1} + \omega^{\alpha}_{jt}, \tag{3a}$$

$$\beta_{jt} = \delta^{\beta}_{j} + \lambda^{\beta}_{j}\beta_{jt-1} + \gamma^{\beta}_{1}\overline{ADV}_{jt-1} + \gamma^{\beta}_{2}\overline{DSC}_{jt-1} + \gamma^{\beta}_{3}\overline{DBR}_{jt-1} + \gamma^{\beta}_{4}\overline{LLN}_{jt-1} + \omega^{\beta}_{jt}. \tag{3b}$$

The γs measure the effect of marketing variables on the base sales and regular price elasticities. These are the central parameters of interest in our analysis as they measure the effect of marketing strategy on brand performance. Standardization of the four marketing variables implies that their parameters estimates are driven by time-varying marketing strategies for a given brand rather than a cross-sectional comparison of marketing strategies across brands. The λ's represent the decay rate of these effects, where λ is positive. A value near 0 implies the effect of marketing strategy is brief whereas a value of 1 implies the effect of the strategy is more enduring (the recursion in 3 implies a geometric decay of marketing effects). We assume all ω's are independently distributed, yet brand specific, with zero mean and a diagonal covariance matrix *W*.

The intuition behind our observation and system equations is that they decompose the short-term from the long-term marketing effects, and the brand effects from the chain effects if necessary. The short-term effects, given by the response parameters in equation (2), capture the contemporaneous effects of marketing variables on a given week's sales of a brand within a given chain. For example, μ_j captures the current period effect of a chain-specific discount on brand sales in a given week. The long-term effects of marketing are captured via the influence of marketing variables on α_{jt} (base sales) and β_{jt} (regular price elasticity) as shown in equations (3a) and (3b). Hence, γ^{α}_2 captures the effect of a brand's cumulative historical discounting on base sales. Likewise, whereas μ_j captures the short-term effect of a *local* or chain-specific discount on sales, ζ_{2j} captures the short-term effect of *national* discounting policy on local or chain level brand sales. One might expect the contemporaneous effect of national discounting to be quite small when controlling for local chain effects because not all stores within a chain adopt the promotion -and this is what we find. In our subsequent elasticity calculations, we focus on change in national brand strategy as opposed to idiosyncratic changes at the chain-level.

Price and marketing mix endogeneity, performance feedback and competitor response. A meta-analysis by Bijmolt, Van Heerde, and Pieters (2005) indicates that price endogeneity plays a major role in price response estimates. To mitigate this bias, we adopt an approach that is analogous to a limited information simultaneous equations approach to the endogeneity problem. As in such models, we replace the supply side model with a linear specification including instrumental variables as the independent variables, and allow for correlation between the demand side error term and the supply side error term. Specifically, we construct the following equation:

$$\begin{aligned}\overline{\ln RPR}_{jst} &= \mu_{0j}^{RPR} + \mu_{1j}^{RPR}\overline{\ln RPR}_{jst-1} + \mu_{2j}^{RPR} TEMP_t \\ &\quad + \mu_{3j}^{RPR}\overline{\ln SALES}_{jt-3}^{Own} + \mu_{4j}^{RPR}\overline{\ln SALES}_{jt-3}^{Cross} + \upsilon_{jst}^{RPR}\end{aligned} \tag{4}$$

The specification assumes that a brand's regular price in a particular chain ($\overline{\ln RPR}_{jst}$) is a manifestation of its (latent) national pricing strategy μ_{0j}^{RPR} . Deviations from this strategy arise from seasonal and random effects. We use lagged regular price ($\overline{\ln RPR}_{jst-1}$) to capture inertia in pricing (Yang, Chen, and Allenby 2003). By including lagged national sales of the focal brand and lagged sum of competing brands' national sales ($\overline{\ln SALES}_{jt-3}^{Own}$ and $\overline{\ln SALES}_{jt-3}^{Cross}$, respectively), this specification also allows us to control for own- and cross-performance feedback.[4] We estimate equations (2) and (4) simultaneously and let error terms, υ_{jst}^{S} and υ_{jst}^{RPR} , be correlated in order to account for price endogeneity in the observation equation. A similar equation is also specified for promotional price ($\overline{\ln PI}_{jst}$).

Finally, we specify an additional equation for each marketing variable to control for performance feedback in the marketing spend. Otherwise, the imputed link between marketing spend and brand performance may be an artifact of the effect that past performance has on marketing spend. Another key advantage of this approach is that it affords a parsimonious control for changes in long-term marketing strategies of *competing brands*, because the sales of these brands are a function of their marketing strategies. Therefore, we include the following regression equation in our system for all four marketing variables:

(5) $$\overline{Z}_{ijt} = \mu_{0j}^{Zi} + \mu_{1j}^{Zi}\overline{Z}_{ijt-1} + \mu_{2j}^{Zi}TEMP_t + \mu_{3j}^{Zi}\overline{\ln SALES}_{jt-3}^{Own} + \mu_{4j}^{Zi}\overline{\ln SALES}_{jt-3}^{Cross} + \upsilon_{ijt}^{Zi},$$

where $\overline{Z}_{ijt}$ is the *i*th marketing variable of brand *j* during week *t*. μ_{1j} captures inertia in marketing and μ_{2j} accounts for seasonality, while the parameters μ_{3j} and μ_{4j} capture, respectively, the own- and cross-performance feedback effects for marketing variable *i*. This specification builds on the work of Horvath et

[4] In the pricing and marketing mix equations –discussed subsequently- we entertained two sets of exogenous variables: (1) lagged dependent variable, lagged own and competitor sales, and (2) these variables in addition to a lagged composite index of competing brands' prices and marketing mix variables (using a sales weighted average to construct this index). Wu-Hausmann tests performed on a static version of our model indicate endogeneity for both sets of variables. Using Sargan's overidentifying restriction test (similar to Basman's J), we find that the null hypothesis that all instruments are exogenous is rejected with both sets of the instruments. We find that omitting the competitive indices and substituting the third lag for sales leads to a set of instruments where one can not reject exogeneity.

al. (2005), who show that own- and cross-performance feedback are more informative than direct competitive action in the prediction of marketing mix activity. In support of this, recent research observes that cross-instrument competitive reactions are predominantly zero (e.g., Ataman, Mela, and Van Heerde 2008; Pauwels 2007; Steenkamp et al. 2005). Equation (5) implies that marketing spend is affected by a geometrically weighted sum of own- and competing-brand sales from the preceding periods. The model therefore captures phenomena such as retailers' disadoption of brands whose sales have been declining for several months (e.g., Franses et al. 1998). Finally, the model also accommodates dynamic dependencies among all the marketing variables via the mediating impact of sales (e.g., Bronnenberg, Mahajan, and Vanhonacker 2000).

Using MCMC techniques we estimate equation (5) together with equations (2) and (4) and let error terms υ_{jst}^{S}, υ_{jst}^{RPR}, υ_{jst}^{PI}, and υ_{ijt}^{Zi} be correlated. Allowing for contemporaneous correlation between sales, pricing and marketing mix equations helps to (i) account for common unobserved shocks that may jointly influence sales and marketing, (ii) control for simultaneity without inducing a causal ordering among the contemporaneous effects, and (iii) capture covariation in marketing expenditures that may arise from retailer category management practices. The details of the estimation procedure are provided in the online appendix.

Note that some of the parameters in equations (2)-(5) are specified as non-time varying. The state space enlarges exponentially with additional time varying parameters and we found the model to yield poor reliability and convergence when all parameters, including those for control variables in equations (2) and (4), and all parameters in equation (5), were allowed to vary. Though the resulting degrees of freedom in Bayesian DLM models are difficult to assess and data dependent due to the precision of the likelihood and priors, it is evident that strong and perhaps unpalatable assumptions would be necessary to identify time-varying parameters for all the regressors.

EMPIRICAL ANALYSIS

We use a novel data set provided by Information Resources Inc. (France) to calibrate our model. These data include five years (1/1/1999 to 1/1/2004) of weekly SKU-store level scanner data for 25 product categories sold in a national sample of 560 outlets representing 21 chains. The 25 categories are chosen to vary across dimensions such as food / nonfood, storable / non-storable, new / mature, etc. In addition, TNS Media Intelligence (France) provided the matching monthly brand-level advertising data. Accordingly, the data includes temporal and cross-sectional changes in (i) advertising strategies, (ii) product offerings, (iii) distribution coverage, and (iv) pricing strategies. One reason we selected France over the United States is that it does not suffer from measurement problems induced by Wal-Mart. Given Wal-Mart sales are growing and as IRI and AC Nielsen do not cover this chain, parameter paths could reflect these changes.

The data's long duration, coverage of the entire mix and manifold categories make the data well suited to address our core research questions. On the other hand, its massive size renders estimation of an SKU-store level model specification infeasible. As such, the data are aggregated to the brand-chain level. We aggregate to the brand level as our central interest pertains to the effect of marketing strategy on brand sales and we aggregate to the chain-level as pricing and other marketing policies tend to be fairly consistent within chains in our data.

Data are aggregated from the SKU-store level to brand-chain level following the procedures outlined in Christen et al. (1997) to avoid any biases due to aggregation. We limit our analyses to the top four chains (184 stores), which account for approximately 75% of the total turnover across all categories, and to three top-selling national brands per category.[5] However there are three categories – dominated by private labels – in which we observe less than three national brands being sold in the top four chains over the entire sample period. This leaves us with 70 national brands in total. The total market share of the top three national brands ranges between 26.1% (Oil) and 79.1% (Carbonated Soft Drinks). We present the

[5] We omit store brands because they do not advertise and their distribution is limited, so we can not use these to contrast elements of the marketing mix.

complete list of variables and their operationalizations in the Appendix. In Table 3 we show the descriptive statistics of our data. Obviously, there is more week-to-week variation in the advertising and discounting variables than in the distribution and product variables. However, because our data spans a long time period (5 years of weekly data), there is sufficient variation in the product and distribution variables to measure their effects.

[TABLE 3 ABOUT HERE]

RESULTS

In this section we first discuss results of the short-term sales model. Next we detail long-term effects including i) the effect of the marketing mix on base sales and regular price elasticities and ii) inertia and performance feedback arising from the marketing expenditures model. We conclude by integrating the long- and short-term models to derive insights regarding the overall effect of marketing strategy on sales over the long- and short-term and across the mix.

The Short-term Effects

We consider three sets of parameters in the sales model (equation 2) for each of the 70 brands; (i) the control variable parameters such as promotional price elasticity, feature/display multiplier, cross-price elasticities, and seasonality parameters, (ii) parameters pertaining to short-term marketing effects, and (iii) the time varying parameters (the intercepts and elasticities).

The promotional price elasticity is –3.35 (reported in Table 4), consistent with Bijmolt, Van Heerde, and Pieters (2005). The mean of the feature and display multipliers, obtained by taking the anti-log transformation, is 1.12, which is comparable to other results in the literature (Van Heerde, Mela, and Manchanda 2004). The regular and promotional cross-price elasticity estimates average .07 and .18, respectively, across all brands, which is also similar to other results in the literature (Sethuraman, Srinivasan, and Kim 1999). The coefficient of average weekly temperature is significant (95% posterior density interval excludes zero) in product categories where sales is expected to exhibit a seasonal pattern (i.e., reaching a peak during summer months in categories like ice cream and carbonated soft drinks, and during winter months in categories like soup and coffee), and insignificant in others.

Table 4 also indicates that all marketing mix variables, on average, have a positive short-term effect on sales. The strongest effects pertain to distribution breadth (.016) and product line length (.015), followed by advertising (.008) and discounting (.0001).

The average regular price elasticity over time and across brands is –1.45, consistent with the results of a recent meta-analysis by Bijmolt, Van Heerde, and Pieters (2005). However, our greater interest lies in how these change over time, to which we turn next.

[TABLE 4, TABLE 5, AND TABLE 6 ABOUT HERE]

The Long-term Effects

Base Sales and Price Elasticity. Of central interest is the long-term effect of marketing strategy on base sales and regular price elasticity (equations 3a and 3b). Across all categories, the marketing effects on base sales and regular price elasticity are given by γ^{α} and γ^{β} respectively (reported in Table 5). For each individual category Table 6 displays the median long-term effect across the brands in the category. For each brand *j* these are given by $\gamma^{\alpha}/(1-\lambda^{\alpha}_{j})$ and $\gamma^{\beta}/(1-\lambda^{\beta}_{j})$.

Table 5 indicates that advertising spending and product line length increase base sales, as expected. The negative effect of discounting reflects that excessive discounting lowers base sales – consistent with deal-to-deal buying patterns. The effect of distribution breadth on *base sales* is negligible as the 90% posterior density interval includes zero. However, as we discuss below, this does not mean that distribution breadth has negligible impact on *sales*.

Table 5 further indicates that product line length and advertising increase regular price elasticity (i.e., make it less negative). The result supports the notion that offering more alternatives and high advertising support help brands better match consumer needs to products and differentiate themselves from the competition. On the other hand, discounting decreases price elasticity (i.e., makes it more negative). This result is consistent with previous research, which suggests that discounts make demand more price elastic (Kopalle, Mela, and Marsh 1999). Finally, Table 5 indicates that the effect of distribution breadth on price elasticity is negative, however the effect can be considered negligible as the

90% posterior density interval includes zero.[6]

Table 6 shows that the magnitudes of the long-term effects on base sales and elasticities vary considerably across categories. Moreover, categories for which the effects on base sales are relatively strong (e.g., diapers, soup) do not necessarily coincide with categories for which the effects on elasticity are relatively strong (e.g., detergent, bath products). This lends support for our two-faceted measures of brand performance. This may be related to the purchase cycle of some of these categories as long-term effects tend to be more enduring as these purchase cycles lengthen –and we overview these duration effects next.

Duration of Base Sales and Price Elasticity Dynamics. Also of interest is the duration of these effects, parameterized by λ in our model (equations 3a and 3b). Given that a brand has done well, one might wonder how long positive effects linger. Conversely, given a brand has done poorly the question is how long it takes to resuscitate it. Across the 70 brands, the intercept decay parameters range between .48 (25^{th} percentile) and .92 (75^{th} percentile), with a median of .69. This implies that 90% duration interval of marketing activity (Leone 1995) range from 3.2 to 28.3 weeks with a median of 6.2 weeks. The median decay for regular price elasticity is .44, ranging between .25 (25^{th} percentile) and .73 (75^{th} percentile), and the implied 90% duration interval range from 1.7 to 7.2 weeks with a median of 2.8 weeks. This implies that the adjustment in elasticity is slightly faster than the adjustment in base sales. In seven categories the effects of the marketing mix on base sales or elasticities appear to be persistent (non-stationary) since the posterior density intervals for decay parameters include 1 (Dekimpe and Hannsens 1999). Overall these dynamics imply it is generally possible to recover from a weak position within a couple of months. However, in some instances it can take ½ year or more to resuscitate a brand.

Price and Marketing Mix Expenditure Dynamics. We summarize the findings that pertain to the regular and promotional price equations (equation 4) and the four marketing mix models (equation 5). First, we compared the fit of a model with no endogeneity and performance feedback to that of a model

[6] We also considered (i) a product variety measure, (ii) a distribution depth variable (analogous to shelf facings), and (iii) a feature/display variable as long-term drivers of brand performance. However, all variables evidenced minimal explanatory power.

with these controls. A log Bayes Factor (West and Harrison 1997) of 12,992 suggests it is critical to control for endogeneity and performance feedback. Second, Table 4 shows that inertia in prices and marketing mix ranges between .62 (distribution) and .92 (line length). Third, better historical performance leads to greater marketing spend (i.e., increased distribution coverage and longer product lines), highlighting the importance of controlling for performance feedback when evaluating the long-term effect of marketing strategy. Finally, we find that cross-sales performance feedback is usually zero.

The Short- and Long-term Effects of Marketing Variables on Sales

So far, our discussion about the long-term effect of marketing variables on *base sales* and *elasticities* has focused on the γs in equation (3). However, to quantify the full impact of marketing variables on *sales*, we also need to consider the direct (contemporaneous) effects of marketing variables on sales via equation (2), the indirect effects via the inertia and feedback effects present in equation (5), and their implications on chain level regular prices and price indices via equation (4). To calculate the full effects of the marketing variables ($\overline{ADV}_{jt}, \overline{DSC}_{jt}, \overline{DBR}_{jt}, \overline{LLN}_{jt}$) on sales, over the short- and long-term, we set each variable at its mean, and then increase each marketing variable, in turn, by 1% in week *t*. The effect on ln(sales) in week *t* (via equation (2)) is the short-term elasticity, $\hat{\eta}_k^s$, where k denotes the element of the mix (e.g., advertising) and s indicates short-term. This shock in marketing also carries forward to future periods in several ways -including inertia (the μ_{1j}^{Zi} in equation 5), performance feedback (the μ_{3j}^{Zi} in equation 5), and the long-term effect on base sales and price elasticity in equation 3. We compute the cumulative implication of this shock for ln(sales) over a time window of 52 weeks (weeks *t*+1, ..., *t*+52), representing the long-term elasticity, $\hat{\eta}_k^l$. The total effect ($\hat{\eta}_k^t$) is given by the long-term effect plus the short-term effect. To compute the relative effect, we calculate $\left|\hat{\eta}_k^p\right| / \sum_k \left|\hat{\eta}_k^p\right|$ where p = {s, l, t}. Table 7 shows the contemporaneous, long-term, and total brand sales elasticity of the marketing mix and Figure 3 presents a pie chart of the relative effects.

[TABLE 7 AND FIGURE 3 ABOUT HERE]

Several striking results emerge. First, the short-term elasticities ($\hat{\eta}_k^s$) of distribution and product are predominant. The distribution elasticity is .13, the product elasticity is .08, the discount elasticity is .06, and the advertising elasticity is .01.[7] The short-term depth elasticity is slightly larger than the mean of .02 reported in Jedidi et al. (1999), while the short-term advertising elasticities appear to be somewhat smaller than the average of .05 for mature brands reported in Lodish et al. (1995).

Second, the long-term elasticities ($\hat{\eta}_k^l$) of product (1.29) and distribution (.61) dwarf the elasticities for advertising (.12) and discounting (-.02). The long-term advertising elasticity (.12) is lower than the empirical generalization (.20) reported in Hanssens, Parson, and Schultz (2001, p. 329). This difference, as well as the ones discussed in the previous paragraph, might be attributable to (a) our inclusion of the full marketing mix as regressors (most studies so far only included a subset (see Table 1), possibly suffering omitted variable biases), or (b) changes in the effectiveness of advertising and promotion over time.

Third, we find that the magnitude of the negative long-term effect of promotion is about one-third of the magnitude of the positive short-term effect, consistent with the result for a single category reported by Jedidi et al. (1999). In contrast, the ratio is reversed for other marketing mix instruments, making the greater long-term impact on brand building evident. For these other instruments, the long-term effects are 4 to 16 times the short-term effects. The larger long-term effect results from an interaction between a large short-term effect and substantial carryover. In particular, product has the highest inertia and distribution has the highest sales performance feedback. As a result the total effect of these instruments is much larger than for promotion.

Last, the total (short-term plus long-term) elasticity $\hat{\eta}_k^t$ (and its share of the sum of total elasticities) of product is 1.37 (60%), while the long-term elasticity for distribution breadth is .74 (32%).

[7] The discount depth elasticity (.06) should not be confused with the price promotion elasticity (-3.35). Note that a 1% increase in discount depth at the chain level (DSC_{jst}), arising from a 1% increase in national discount depth (DSC_{jt}), corresponds to a much smaller decrease in the price index variable (PI_{jst} = 1-DSC_{jst}). This relationship coupled with the low average discount depth observed in our data (1.8%) explains the modest magnitude of discount depth elasticity.

In sharp contrast, the effect of advertising is only .13 (6%) and for discounts it is .04 (2%). Hence we find evidence that distribution and product play the major role in discriminating between the performance of mature brands in spite of the emphasis of prior research on discounts and advertising (e.g., Jedidi et al. 1999). This result is consistent with the common wisdom that distribution and product are among the most important components of marketing strategy.

SUMMARY

While marketing managers spend many billions of dollars annually on their marketing programs, few studies systematically assess the long-term effect of these programs over many brands and categories. Moreover, extant research focuses largely on advertising and promotions (see Table 1), but not on product or distribution.

This study attempts to address both the data and the modeling requirements. We use five years of weekly data across 25 categories, 70 brands sold in the four largest chains in France. By relating the performance of these brands to their entire marketing mix strategy, we afford insights into which strategies are most likely to lead to long-term advantages for brands. We apply our data to a Dynamic Linear Model (DLM), which allows us to model both sales and the marketing mix as dependent variables, and helps us to accommodate endogeneity, performance feedback, and competitive interactions (via cross performance feedback effects).

Using the DLM, we link marketing strategy to two components of brand performance, base sales and regular price elasticity. We find:

- All aspects of the marketing mix exhibit a positive short term direct effect on sales -most notably distribution and line length.
- The mix also evidences indirect effects through base sales and price response. Base sales are positively affected by advertising but negatively affected by discounting over the long run. Regular price elasticities are decreased by discounting and distribution, but they are increased by advertising and line length. We suspect the negative effect of distribution on price elasticity is due to increased potential for consumers to shop across stores.

- The median 90% average decay of the mix effect on base sales is around 6.2 weeks. The corresponding figure for elasticities is 2.8 weeks.
- Dynamics are also present via performance feedback and inertia in spending. Performance feedback is strongest for distribution while inertia is strongest for product.
- When combined, all of these effects indicate that product (60%) and distribution (32%) have a substantially larger relative effect on brand sales over the long run than discounting (2%) or advertising (6%).
- Moreover, we find that the magnitude of the dynamic effect of a promotion is one-third of the magnitude of its contemporaneous effect. This ratio is reversed for other aspects of the marketing mix, suggesting their greater potential to make an enduring impact on brand sales.

LIMITATIONS

These findings are subject to several notable limitations, some of which point out several future research opportunities:

- The DLM is well-suited to linking marketing activity to intercepts and elasticities but can not easily be scaled to a large number of variables, periods and observations because a) the state-space explodes and, along with it, the computer memory needed for estimation, and b) convergence of each model run takes weeks. Therefore our use of the DLM amplifies the trade-off between model parsimony and completeness. Accordingly, we have made a number of assumptions to render the analysis feasible, including:
    - Our model does not allow for different decay factors for different marketing variables. One can write a canonical transfer function DLM to overcome this limitation and estimate different decay parameters for each marketing variable using a data augmentation step in the Gibbs sampler.
    - Several potential interactions exist in the marketing mix. For example, advertising itself may facilitate new distribution. We control for these effects indirectly via lagged performance

feedback, which embed the marketing actions pursued by firms in preceding periods.

- We presume the effects of feature and display are fixed over time. Undoubtedly, these effects can change over time with marketing strategy. Expansion of the model to accommodate these effects would render such insights unreliable as a result of increased model complexity. In an analysis not reported herein, we estimate a simpler version of the DLM wherein all parameters are time varying but the time paths are not specified to vary with the marketing mix. The estimated parameter paths for price and the intercept are largely the same as observed in our model, suggesting that the omission of time varying effects for feature and display does not bias our results.
- We aggregate data to the chain level. It would be desirable to extend this research to the store level, as that would allow us to study inter-retail price competition. Chain-level measures are more noisy and the reduction in observations reduces power. As a result, our research is a conservative test of our hypotheses. Chain-level analysis is not uncommon in marketing (e.g., Slotegraaf and Pauwels 2008; Srinivasan et al. 2004), perhaps because marketing activity tends to be correlated across stores within a chain.
- We consider the top 4 chains and the largest 3 brands in each category. As such, our results should be interpreted from the perspective of managers with large brands selling through predominantly large chains. It would be interesting to consider whether the results generalize to smaller niche brands and outlets (e.g., Ataman, Mela, and van Heerde 2008; Slotegraaf and Pauwels, 2008).

Most of these extensions are tangential to our central research objectives. Yet, we believe they would form the basis for future work to further enhance our comprehension of the role marketing plays on performance in the long-term.

- Due to lack of data we cannot include the perceived quality of the brands (e.g., Aaker and Jacobson 1994) as a driver of brand performance. Perceived quality is a fixed effect, however, so our standardization should control for its omission.

- We consider retail price elasticities when evaluating the effect of observed marketing strategies on brand performance. However, retail prices embed both the behaviors of retailers and the firms that supply them. Accordingly, a formal accounting of the role of retailers in driving brand price elasticities would help firms disentangle those aspects of marketing strategy more salient to the firm and those more relevant to the retailer.

Despite these limitations, we believe our paper makes an important first step in documenting the overall long-term effects of the entire marketing mix on brand sales. We hope this study will stimulate additional research that analyzes these effects in more detail, enabling even more finely-tuned recommendations for marketing executives.

REFERENCES

Aaker, David A. (1991), *Managing Brand Equity*, Free Press, New York.

Aaker, David A. and Robert Jacobson (1994), "The Financial Information Content of Perceived Quality," *Journal of Marketing Research*, 31 (May), 191-201.

Ailawadi, Kusum L., Karen Gedenk, Christian Lutzky, and Scott A. Neslin (2007), "Decomposition of the Sales Impact of Promotion-Induced Stockpiling," *Journal of Marketing Research* 44, (3), 450-467.

Ailawadi, Kusum L., Donald R. Lehmann, and Scott A. Neslin (2003), "Revenue Premium as an Outcome Measure of Brand Equity," *Journal of Marketing*, 67 (October), 1–17.

Ataman, M. Berk, Carl F. Mela, and Harald J. Van Heerde (2008), "Building Brands," *Marketing Science*, 27 (6), 1036-1054.

Baghestani, Hamid (1991), "Cointegration Analysis of the Advertising-Sales Relationship," *The Journal of Industrial Economics*, 39 (December), 671-81.

Bass, Frank M., Norris Bruce, Sumit Majumdar, B.P.S. Murthi (2007), "Wearout Effects of Different Advertising Themes: A Dynamic Bayesian Model of Advertising-Sales Relationship," *Marketing Science*, 26(2), 179-195.

Berger, Jonah, Michaela Draganska, and Itamar Simonson (2007), "The Influence of Product Variety on Brand Perception and Choice," *Marketing Science*, 26 (July-August), 460-472.

Bijmolt, Tammo H.A., Harald J. Van Heerde, and Rik G.M. Pieters (2005), "New Empirical Generalizations on the Determinants of Price Elasticity," *Journal of Marketing Research*, 42 (May), 141-156.

Boulding, William, Eunkyu Lee, and Richard Staelin (1994), "Mastering the Mix: Do Advertising, Promotion and Sales Force Activities Lead to Differentiation?," *Journal of Marketing Research*, 31 (May), 159–172.

Bronnenberg, Bart J., Vijay Mahajan, and Wilfried R. Vanhonacker (2000), "The Emergence of Market Structure in New Repeat-Purchase Categories: The Interplay of Market Share and Retailer Distribution," *Journal of Marketing Research*, 37 (February), 16–31.

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

Carter, C. and R. Kohn (1994), "On Gibbs Sampling for State Space Models," *Biometrika*, 81 (3), 541–53.

Chib, Siddhartha and Edward Greenberg (1995), "Understanding the Metropolis-Hastings Algorithm," *American Statistician*, 49 (4) 327-335.

Christen, Markus, Sachin Gupta, John C Porter, Richard Staelin, and Dick R Wittink (1997), "Using Market-level Data to Understand Promotion Effects in a Nonlinear Model," *Journal of Marketing Research*, 34 (August), 322–334.

Clarke, Darral G. (1976), “Econometric Measurement of the Duration of Advertising Effect on Sales,” *Journal of Marketing Research*, 13 (November), 345–57.

Dekimpe, Marnik G. and Dominique M. Hanssens (1995), “The Persistence of Marketing Effects on Sales,” *Marketing Science*, 14 (1), 1-21.

Dekimpe, Marnik G. and Dominique M. Hanssens (1999), “Sustained Spending and Persistent Response: A New Look at Long-term Marketing Profitability,” *Journal of Marketing Research*, 39 (November), 397–412.

Dekimpe, Marnik G., Dominique M. Hanssens, and Jorge M. Silva-Risso (1999), “Long-run Effects of Price Promotions in Scanner Markets,” *Journal of Econometrics*, 89 (March/April), 269–291.

Foekens, Eijte W., Peter S.H. Leeflang, and Dick R. Wittink (1999), “Varying Parameter Models to Accommodate Dynamic Promotion Effects,” *Journal of Econometrics*, 89 (March/April), 249–268.

Fox, Edward J., and Stephen J. Hoch (2005), “Cherry-Picking,” *Journal of Marketing*, 69 (1), 46-62.

Franses, Philip Hans, Teun Kloek and A. Lucas (1998), “Outlier Robust Analysis of Long-run Marketing Effects for Weekly Scanning Data,” *Journal of Econometrics*, 89 (1-2), 293-315.

Frühwirth-Schnatter, Sylvia (1994), “Data Augmentation and Dynamic Linear Models,” *Journal of Time Series Analysis*, 15 (2), 183–202.

Golder, Peter N. and Gerard J. Tellis (1997), "Will It Ever Fly? Modeling The Growth of New Consumer Durables," *Marketing Science*, 16 (3), 256-270.

Hanssens, Dominique M., Leonard J. Parsons and Randall L. Schultz (2001), *Market Response Models: Econometric and Time Series Analysis*, 2nd edition, Boston: Kluwer Academic Publishers.

Horvath, Csilla, Peter S.H. Leeflang, Jaap E. Wieringa, and Dick R. Wittink (2005), “Competitive reaction- and feedback effects based on VARX models of pooled store data,” *International Journal of Research in Marketing*, 22 (December), 415–426.

Jedidi, Kamel, Carl F. Mela, and Sunil Gupta (1999), “Managing Advertising and Promotion for Long-run Profitability,” *Marketing Science*, 18 (1), 1–22.

Kamakura, Wagner A. and Gary J. Russell (1993), “Measuring Brand Value with Scanner Data,” *International Journal of Research in Marketing*, 10 (March), 9–22.

Kaul, Anil, and Dick R. Wittink (1995). “Empirical Generalizations About the Impact of Advertising on Price Sensitivity and Price,” *Marketing Science*, 14 (3 part 2 of 2), G151-G160.

Keane, Michael (1997), “Modeling Heterogeneity and State Dependence in Consumer Choice Behavior,” *Journal of Business and Economic Statistics*, 15 (3), 310-327.

Keller, Kevin L. (1993), “Conceptualizing, Measuring, and Managing Customer-based Brand Equity,” *Journal of Marketing* 57, 1–22.

Kirmani, Amna and Peter Wright (1989), “Money Talks: Perceived Advertising Expenditures and Expected Product Quality,” *Journal of Consumer Research*, 16 (December), 344–353.

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

Leone, Robert P. (1995), “Generalizing What is Known of Temporal Aggregation and Advertising Carryover,” *Marketing Science*, 14 (3 part 2 of 2), G141-G150.

Lodish, Leonard M., Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, and Mary Ellen Stevens (1995), “How T.V. Advertising Works: A Meta-Analysis of 389 Real World Split Cable T.V. Advertising Experiments,” *Journal of Marketing Research*, 32 (May), 125–139.

Mela, Carl F., Sunil Gupta, and Donald R. Lehmann (1997), “The Long-term Impact of Promotion and Advertising on Consumer Brand Choice,” *Journal of Marketing Research*, 34 (May), 248–261.

Mela, Carl F., Sunil Gupta, and Kamel Jedidi (1998), “Assessing Long-term Promotional Influences on Market Structure,” *International Journal of Research in Marketing*, 15 (May), 89–107.

Mela, Carl F., Kamel Jedidi, and Douglas Bowman (1998), “The Long-term Impact of Promotions on Consumer Stockpiling,” *Journal of Marketing Research*, 35 (May), 250–262.

Mitra, Anusree and John G. Lynch (1995), “Toward a Reconciliation of Market Power and Information Theories of Advertising Effects on Price Elasticity,” *Journal of Consumer Research*, 21 (March), 644–659.

MSI Research Priorities (2008). Cambridge, MA: Marketing Science Institute.

Naik, Prasad A., Murali Mantrala, and Alan Sawyer (1998), “Planning Media Schedules in the Presence of Dynamic Advertising Quality,” *Marketing Science*, 17 (3), 214-235.

Nicholson, Walter (1972), *Microeconomic Theory: Basic Principles and Extensions*, Hinsadale, IL: The Dryden Press.

Nijs, Vincent R., Marnik G. Dekimpe, Jan-Benedict E. M. Steenkamp, and Dominique M. Hanssens (2001), “The Category-demand Effects of Price Promotions,” *Marketing Science*, 20 (1), 1–22.

Papatla, Purushottam and Lakshman Krishnamurthi (1996), “Measuring the Dynamic Effects of Promotions on Brand Choice,” *Journal of Marketing Research*, 33 (February), 20–35.

Pauwels, Koen (2004), “How Dynamic Consumer Response, Competitor Response, Company Support, and Company Inertia Shape Long-Term Marketing Effectiveness,” *Marketing Science*, 23 (3), 596-610.

Pauwels, Koen (2007), “How Retailer and Competitor Decisions Drive the Long-Term Effectivness of Manufacturer Promotions for Fast Moving Consumer Goods,” *Journal of Retailing*, 83 (3), 297-308.

Pauwels, Koen and Dominique M. Hansens (2007), "Performance Regimes and Marketing Policy Shifts," *Marketing Science*, 26(3), 293-311.

Pauwels, Koen, Jorge Silva-Risso, Shuba Srinivasan, and Dominique M Hanssens (2004), “New Products, Sales Promotions, and Firm Value: The Case of the Automobile Industry,” *Journal of Marketing*, 68 (October), 142-56.

Pauwels, Koen, Dominique M. Hanssens, and S. Siddarth (2002), “The Long-term Effects of Price

Promotions on Category Incidence, Brand choice, and Purchase Quantity," *Journal of Marketing Research*, 39 (November), 421–439.

Sethuraman, Raj, V. Srinivasan, and Doyle Kim (1999), "Asymmetric and Neighborhood Cross-Price Effects: Some Empirical Generalizations," *Marketing Science*, 18 (1), 23–41.

Simon, Carol J, Mary W. Sullivan (1993), "The Measurement and Determinants of Brand Equity: A Financial Approach," *Marketing Science*, 12 (1), 28-52.

Slotegraaf, Rebecca J. and Koen Pauwels (2008), "The Impact of Brand Equity and Innovation on the Long-Term Effectiveness of Promotions," *Journal of Marketing Research*, 45 (June), 293-306.

Srinivasan, Shuba, Peter TL Popkowski Leszczyc, and Frank M. Bass (2000), "Market Share Response and Competitive Interaction: The Impact of Temporary, Evolving and Structural Changes in Prices," *International Journal of Research in Marketing*, 17 (December), 281-305

Srinivasan, Shuba, Koen Pauwels, Dominique M. Hanssens, and Marnik G. Dekimpe (2004), "Do Promotions Benefit Manufacturers, Retailers, or Both?," *Management Science*, 50 (May), 617-29

Sriram, S., Subramanian Balachander, and Manohar U. Kalwani (2007), "Monitoring the Dynamics of Brand Equity Using Store-Level Data," *Journal of Marketing*, 71 (April), 61-78.

Steenkamp, Jan-Benedict E.M., Vincent Nijs, Dominique M. Hanssens, and Marnik G. Dekimpe (2005), "Competitive Reactions to Advertising and Promotion Attacks," *Marketing Science* 24 (1), 35-54.

Swait, Joffre, Tülin Erdem, Jordan Louviere, and Chris Dubelaar (1993), "The Equalization Price: A Measure of Consumer-perceived Brand Equity," *International Journal of Research in Marketing*, 10 (March), 23–45.

Tellis, Gerard J. and Philip Hans Franses (2006), "Optimal Data Interval for Estimating Advertising Response," *Marketing Science*, 25 (3), 217-229.

Van Heerde, Harald J., Peter S.H. Leeflang, and Dick R. Wittink (2000), "The Estimation of Pre- and Postpromotion Dips with Store-Level Scanner Data," *Journal of Marketing Research*, 37 (August), 383–395.

Van Heerde, Harald J., Carl F. Mela, and Puneet Manchanda (2004), "The Dynamic Effect of Innovation on Market Structure," *Journal of Marketing Research*, 41 (May), 166-183.

West, Mike and Jeff Harrison (1997), *Bayesian Forecasting and Dynamic Models*, 2nd edn, Springer-Verlag, New York.

Yang, Sha, Yuxin Chen, and Greg Allenby (2003), "Bayesian Analysis of Simultaneous Demand and Supply," *Quantitative Marketing and Economics*, 1 (3), 251-275.

Yoo, Boonghee, Naveen Donthu, and Sungho Lee (2000), "An Examination of Selected Marketing Mix Elements and Brand Equity," *Journal of Academy of Marketing Science*, 28 (Spring), 195–211.

Table 1: Current Literature on Long-term Effects of Marketing Variables

	Effect of				Effect on	Modeling Approach	# Categories
	Promotion	Advertising	Distribution	Product			
Clarke (1976)		✓			Brand Sales	VPM	1
Baghestani (1991)		✓			Brand Sales	VAR	1
Dekimpe and Hanssens (1995)		✓			Chain Sales	VAR	1
Papatla and Krishnamurthi (1996)	✓				Choice	VPM	1
Mela, Gupta, and Lehmann (1997)	✓	✓			Choice	VPM	1
Mela, Jedidi, and Bowman (1998)	✓				Incidence and Quantity	VPM	1
Mela, Gupta, and Jedidi (1998)	✓	✓			Market Structure	Mixed	1
Kopalle, Mela, and Marsh (1999)	✓				Brand Sales	VPM	1
Jedidi, Mela, and Gupta (1999)	✓	✓			Choice and Quantity	VPM	1
Foekens, Leeflang, and Wittink (1999)	✓				Brand Sales	VPM	1
Dekimpe and Hanssens (1999)	✓	✓			Brand Sales	VAR	1
Dekimpe, Hanssens, and Silva-Risso (1999)	✓				Brand and Cat. Sales	VAR	4
Srinivasan, Leszczyc, and Bass (2000)	✓		✓		Market Share	VAR	2
Bronnenberg, Mahajan, and Vanhonacker (2000)	✓	✓	✓		Market Share	VAR	1
Nijs et al. (2001)	✓				Category Sales	VAR	560
Pauwels, Hanssens, and Siddarth (2002)	✓				Incidence, Choice and Quantity	VAR	2
Srinivasan et al. (2004)	✓				Margin and Revenue	VAR	21
Pauwels (2004)	✓	✓		✓	Brand Sales	VAR	1
Van Heerde, Mela, and Manchanda (2004)				✓	Market Structure	VPM (DLM)	1
Pauwels et al. (2004)	✓			✓	Financial measures	VAR	1
Steenkamp et al. (2005)	✓	✓			Brand Sales	VAR	442
Sriram, Balachander, and Kalwani (2007)	✓	✓		✓	Brand Sales	VPM	2
Ataman, Mela, and Van Heerde (2008)	✓	✓	✓	✓	Brand sales (new brands only)	VPM-SE (DLM)	22
Slotegraaf and Pauwels (2008)	✓			✓	Brand Sales	VAR	7
THIS PAPER	✓	✓	✓	✓	Brand sales and Elasticity	VPM-SE (DLM)	25

Notes: VPM = Varying Parameter Model; VAR = Vector Autoregressive model; DLM = Dynamic Linear Model; SE = System of Equations

Table 2: Expected Marketing Mix Effects on Base Sales and Regular Price Elasticity

Variable	Operationalization	Predicted Effect on	
		Base sales (=intercept)	Regular price elasticity[1]
Discounting	Discount depth	?	Negative
Advertising	Expenditure	Positive	Positive
Distribution	%ACV weighted distribution	Positive	?
Line Length	Number of SKUs	Positive	Positive

[1] A positive effect on regular price elasticity means that the elasticity become less negative; a negative effect means it become more negative.

Table 3: Descriptive Statistics

Category	Number of Brands	Market Share (%)	Discount Depth (%)		Advertising (10^5 EUR)		Distribution (% ACV)		Line Length (# SKUs)	
		Mean*	Mean	Variance**	Mean	Variance	Mean	Variance	Mean	Variance
Bath Products	3	9.9	2.1	1.8	.584	0.9	99.9	0.1	50.8	48.9
Beer	3	17.4	2.1	1.8	1.786	6.5	100.0	---	31.0	12.1
Coffee	3	14.4	2.9	2.3	2.877	5.6	100.0	---	36.6	49.9
Chips	1	32.8	3.4	3.2	---	---	99.8	0.3	46.6	169.4
Cereals	3	25.4	1.7	1.2	3.784	7.1	96.8	9.5	32.9	13.7
Soft Drinks	3	26.4	2.4	1.3	2.825	5.4	99.8	0.8	37.2	28.1
Diapers	3	20.8	1.2	1.0	.835	1.0	99.7	0.8	51.1	548.4
Detergent	3	15.6	1.4	2.1	2.891	2.6	100.0	---	43.5	170.2
Feminine Needs	3	18.9	.9	0.6	1.791	1.4	100.0	---	36.2	30.8
Frozen Pizza	3	15.8	2.4	2.4	.396	1.0	97.2	7.5	14.8	17.5
Ice Cream	3	10.1	3.4	3.9	.664	2.2	98.7	3.5	60.0	739.1
Mayonnaise	3	23.9	1.2	1.3	.818	1.2	99.7	0.2	48.8	50.0
Oil	3	8.7	1.6	1.4	.690	0.9	99.8	0.2	21.2	8.5
Pasta	3	20.7	2.5	1.5	1.126	1.7	100.0	---	105.2	156.5
Paper Towel	1	33.9	2.6	1.8	.782	1.4	99.0	1.2	12.4	5.3
Shaving Cream	3	17.3	1.0	0.7	.123	0.2	99.7	0.7	27.2	36.1
Shampoo	3	11.3	1.5	1.0	1.776	2.1	99.9	0.0	41.3	87.5
Soup	3	24.1	1.0	0.9	1.193	3.5	99.7	0.2	67.2	107.8
Tea	3	17.2	.4	0.2	.282	0.4	96.7	15.5	27.8	7.3
Toothpaste	3	17.2	1.3	1.3	1.304	1.3	99.9	0.1	34.3	44.1
Toilet Tissue	3	14.3	1.9	1.1	.352	0.7	97.3	6.5	17.7	5.3
Window Cleaner	2	29.4	.6	0.4	.027	0.1	98.0	6.0	6.0	1.9
Water	3	10.4	1.2	0.9	2.492	6.6	100.0	---	25.0	10.6
Yogurt Drinks	3	26.3	1.8	3.0	.246	0.4	98.7	7.2	11.3	7.8
Yogurt	3	10.8	1.0	0.7	1.030	3.9	99.6	1.9	26.4	11.9
All Categories	70	18.9	1.8	1.5	1.2268	2.3	99.	2.5	36.49	94.8

[*] Average across all weeks and brands within a category.

[**] Average across all brands within a category.

Table 4: Parameter Estimates of Sales and Marketing Mix Models

Equation	Coefficient	Mean*	Variance*
Sales	Feature/Display	.106	.007
	Price Index (Own)	-3.348	6.136
	Price Index (1st Competitor)	.160	.218
	Price Index (2nd Competitor)	.195	.135
	Regular Price (1st Competitor)	.041	.142
	Regular Price (2nd Competitor)	.091	.102
	Temperature	.001	.000
	Christmas	-.079	.049
	New Year	.012	.035
	Easter	.068	.005
	Ascension	-.021	.002
	Bastille Day	-.015	.001
	Assumption	-.052	.003
	Advertising	.008	.004
	Discounting	.000	.001
	Distribution	.016	.003
	Line Length	.015	.009
Advertising	Constant	-.001	.000
	Temperature	.001	.000
	Lagged Advertising	.851	.002
	Own Performance Feedback	-.028	.004
	Cross Performance Feedback	.004	.004
Discounting	Constant	.000	.000
	Temperature	.002	.000
	Lagged Discounting	.707	.007
	Own Performance Feedback	.000	.011
	Cross Performance Feedback	.003	.008
Distribution	Constant	.005	.000
	Temperature	-.002	.000
	Lagged Distribution	.624	.090
	Own Performance Feedback	.037	.015
	Cross Performance Feedback	-.012	.010
Line Length	Constant	.006	.000
	Temperature	.002	.000
	Lagged Line Length	.923	.004
	Own Performance Feedback	.003	.004
	Cross Performance Feedback	-.005	.005
Regular Price	Constant	.000	.000
	Temperature	.000	.000
	Lagged Regular Price	.898	.003
	Own Performance Feedback	.000	.000
	Cross Performance Feedback	.000	.000
Price Index	Constant	.000	.000
	Temperature	.000	.000
	Lagged Price Index	.703	.007
	Own Performance Feedback	.000	.000
	Cross Performance Feedback	.000	.000

* Mean and variance across median estimates for 70 brands.

Table 5: Marketing Mix Effects on Intercepts and Elasticities

	Expected Effect on Intercept	Elasticity	*Estimated Effect on* Intercept *Median*	[5^{th} *and* 95^{th} *percentile*]	Elasticity *Median*	[5^{th} *and* 95^{th} *percentile*]
Effect of						
Discounting	?	Negative	-.0044	[-.0061,-.0029]	-.0119	[-.0067,-.0177]
Advertising	Positive	Positive	.0069	[.0052,.0086]	.0083	[.0014,.0139]
Distribution	Positive	?	-.0008	[-.0026,.0012]	-.0047[b]	[-.0128,.0027]
Line Length	Positive	Positive	.0012[a]	[-.0001,.0025]	.0123	[.0074,.0182]

a. The effect of line length on base sales crosses zero at 93rd percentile.
b. The effect of distribution breadth on elasticity crosses zero at 86th percentile.

Table 6: Long-term Effects of Marketing Mix Effects on Intercepts and Elasticities*

	Intercept				Elasticity			
Category	Discounting	Advertising	Distribution	Line Length	Discounting	Advertising	Distribution	Line Length
Bath Products	-.010	.016	-.002	.003	-.415	.268	-.161	.431
Beer	-.021	.032	-.004	.005	-.017	.012	-.007	.016
Coffee	-.034	.052	-.006	.009	-.022	.014	-.009	.022
Chips	-.014	.022	-.003	.004	-.021	.014	-.009	.022
Cereals	-.014	.023	-.002	.004	-.022	.015	-.009	.023
Soft Drinks	-.020	.032	-.004	.005	-.016	.011	-.007	.018
Diapers	-.181	.278	-.031	.048	-.022	.015	-.008	.022
Detergent	-.008	.013	-.001	.002	-.035	.023	-.014	.038
Feminine Needs	-.009	.014	-.002	.002	-.023	.016	-.009	.024
Frozen Pizza	-.041	.063	-.007	.011	-.026	.018	-.010	.027
Ice Cream	-.068	.105	-.012	.018	-.076	.049	-.029	.078
Mayonnaise	-.033	.052	-.006	.009	-.017	.011	-.006	.016
Oil	-.014	.021	-.002	.004	-.019	.011	-.007	.018
Pasta	-.010	.015	-.002	.003	-.025	.016	-.011	.025
Paper Towel	-.121	.184	-.020	.031	-.021	.014	-.008	.020
Shaving Cream	-.007	.011	-.001	.002	-.018	.012	-.008	.019
Shampoo	-.011	.017	-.002	.003	-.017	.011	-.006	.017
Soup	-.069	.106	-.012	.018	-.042	.024	-.015	.043
Tea	-.009	.014	-.002	.002	-.015	.011	-.006	.017
Toothpaste	-.008	.013	-.001	.002	-.025	.018	-.010	.026
Toilet Tissue	-.010	.014	-.002	.002	-.029	.020	-.011	.030
Window Cleaner	-.013	.020	-.002	.003	-.031	.020	-.012	.032
Water	-.008	.012	-.001	.002	-.016	.011	-.006	.017
Yogurt Drinks	-.099	.149	-.017	.025	-.066	.042	-.027	.075
Yogurt	-.015	.023	-.002	.004	-.020	.013	-.008	.020
ALL	-.014	.022	-.002	.004	-.022	.015	-.009	.022

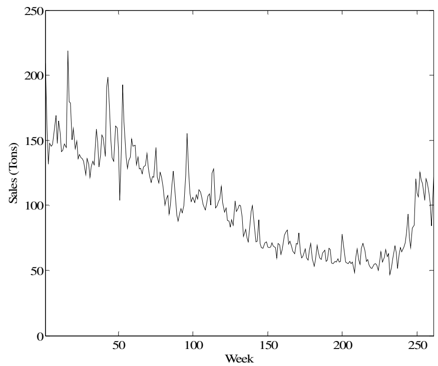
* The long-term effects of marketing variables are computed over a five year horizon. Table entries are medians across brands in a product category.

Table 7: Sales Impact of 1% Temporary Increase in Marketing Support (%)*

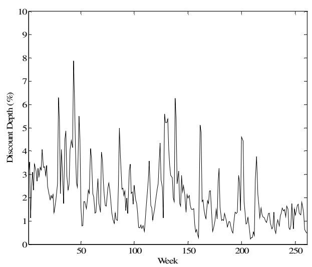
	Contemporaneous	Long-term	Total
Discounting	.06	-.02	.04
Advertising	.01	.12	.13
Distribution	.13	.61	.74
Line length	.08	1.29	1.37

* Table entries are medians across brands.

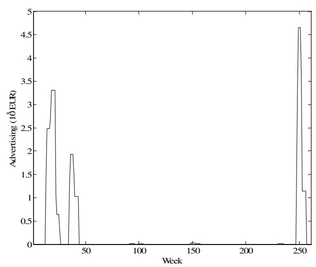
Figure 1: Contraction Case – Brand C



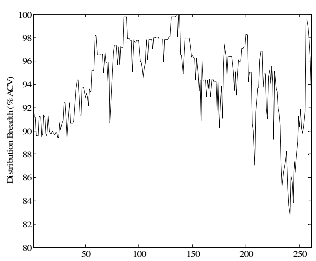
(a) Sales



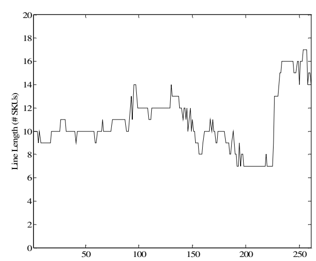
(b) Discount Depth



(c) Advertising Spending

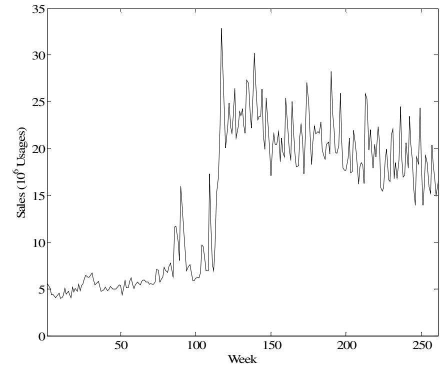


(d) Distribution Breadth

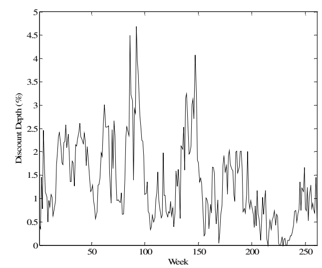


(e) Product Line Length

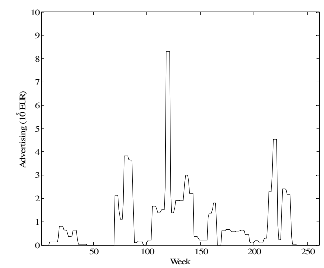
Figure 2: Growth Case – Brand G



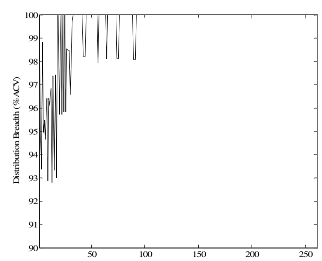
(a) Sales



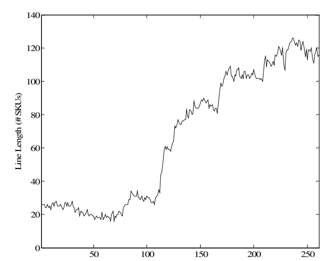
(b) Discount Depth



(c) Advertising Spending

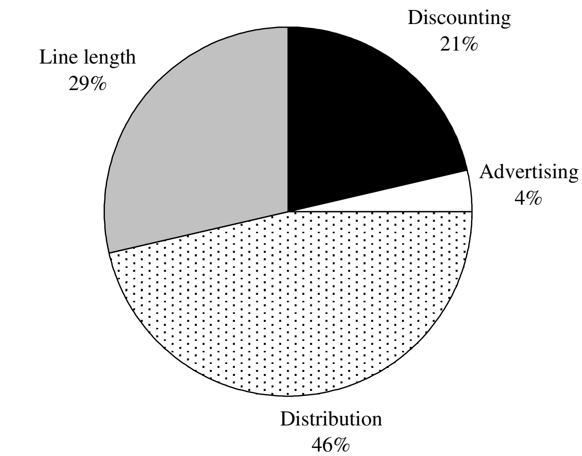


(d) Distribution Breadth

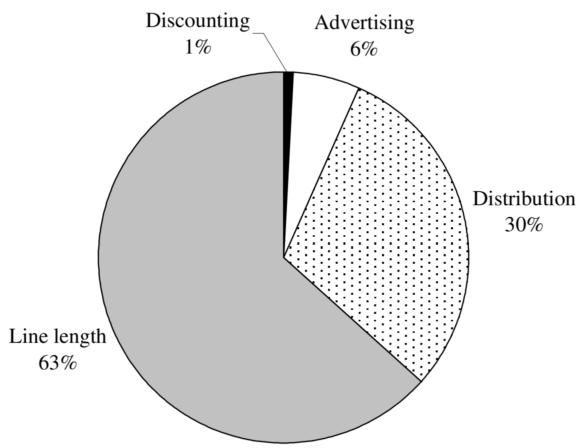


(e) Product Line Length

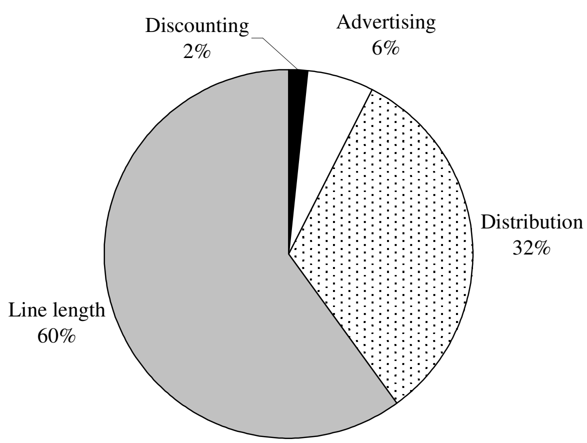
Figure 3: Relative Elasticities across the Mix



(a) Short-term Elasticity



(b) Long-term Elasticity



(c) Total Elasticity

APPENDIX: Variable Operationalizations

Observation (Sales) Equation Variables: The dependent variable of the observation equation is sales volume, calculated as the ACV weighted geometric average of total sales of a brand in a given store-week, across stores in a given chain.[8] The core independent variable is regular price, for which we use the regular price series provided in the IRI data sets. It represents the normal price in the absence of a price discount. We aggregate it similar to Mela, Gupta, and Lehmann (1997), using the minimum regular price per 1000 volume units across SKUs of a brand in a given store and week as the regular price for that brand. This measure has the added benefit of not being sensitive to the particular sales weighting scheme selected. Moreover, it exploits price variation in the data that might be understated in the event one major SKU lowers its regular price. We calculate average chain level brand regular price in the nonlinear fashion outlined by Christen et al. (1997). In addition, we include competing brands' regular prices. We also include own- and cross-brand price index variable (actual price over regular price) to control for promotional price discounts. We assume that a brand is on feature or display if any SKU of that brand is on feature or display in a given store and week. Chain-level feature and display variables are calculated by taking the ACV weighted arithmetic average across stores in a given week (see Christen et al. 1997). The feature and display variable are set to zero when there is a price discount of five percent or more. The benefit of this transformation is a considerable reduction in correlation between price and the variable for feature and display (Van Heerde, Leeflang, and Wittink 2000).[9] As such the feature and display variable measures the effects of these activities in the absence of a price cut, while the price variable measures the impact of price changes that are possibly communicated via feature and display. Finally we use average weekly temperature to account for any seasonal patterns inherent in sales, and include dummy variables to control for Christmas, New Year, Easter, Ascension, Bastille Day, and Assumption.

[8] We use current period store-level ACVs for categories as these are parsimonious to construct and vary negligibly from historical ACVs. Therefore the choice of time frame is immaterial.
[9] We also estimated a model with feature and display not set to zero when there was a price discount. We found the collinearity to be sufficiently large that the price and promotion parameters were not well identified.

State Equation Variables: We operationalize long-term marketing strategies from the following weekly measures. The model then creates a geometric-decay weighted average of the weekly variables to capture their long-term effect (see Equations 3a and 3b). The *price discount fraction* is measured as one minus the ratio of the actual to the regular price. National level averages are calculated across stores and chains in a linear fashion. We construct a weekly *advertising expenditure* variable from our monthly data by dividing the monthly figures by the number of days in a month, and then summing across days for the corresponding weeks (Jedidi, Mela, and Gupta 1999).[10] Following Bronnenberg, Mahajan, and Vanhonacker (2000), we use ACV weighted distribution as a measure of *breadth of distribution*. ACV weights a product's distribution by the total dollar volume sold through a particular store. Thus, ACV gives more distribution credit for an item that is carried in a large dollar volume store than it does for a small dollar volume store. Finally, we use *line length* as the product variable: the number of products that is available for a given brand in a given chain in a given time period (week).

[10] Tellis and Franses (2005) indicate that data interval bias exists when estimating at the higher levels of time aggregation. In contrast, we use the lowest level of time aggregation

WEB APPENDIX

Appendix A: A Primer on Dynamic Linear Models

During the last decade the marketing literature has shown a growing interest in measuring the long-term effects of marketing activity on brand performance. This increased attention led to the development and/or application of various time series analysis techniques in marketing. These modeling techniques can be grouped under two *seemingly* alternative approaches: (i) VAR (vector autoregressive) and VEC (vector error correction) models (e.g., Pauwels, Hanssens and Siddarth 2002; Fok et al. 2006), and (ii) Dynamic Linear Models (e.g., Van Heerde, Mela, and Manchanda 2004). The basic idea of VAR and VEC models is to specify the dependent variable (e.g., a vector with the sales for the top five brands in a category) as a function of current and lagged values of marketing instruments and lagged values of sales. The long-term effects of the marketing variables on sales are measured by using the estimated model to calculate what-if scenarios ("Impulse response functions"). For example, what happens with sales if we increase advertising by X percent?

Dynamic Linear Models (DLMs) can also capture the long-term effects of marketing variables. It consists of an observation equation and a state equation. The observation equation looks like a standard regression model, albeit with time-varying intercept and response parameters. For example, the dependent variable in an observation equation is a vector of sales of the brands in a category, and their marketing instruments are the independent variables. The state equation describes how the regression parameters (intercept and response parameters) evolve over time, possibly under the influence of other variables such as past advertising. The core difference with VAR and VEC models is that the long-term effects in DLM are mediated by changes in model parameters in the state equation. For example, while a VAR or VEC model captures the direct effect of past advertising on sales, in a DLM, the effect of past advertising on sales goes via the time-varying intercept of the sales model. This time-varying intercept is part of the state equation, and it may be modeled as a function of its own past value, past advertising and other marketing instruments. In that sense, the state equation represents an extra layer in DLMs that is not present in VAR or VEC models.

In fact these modeling traditions are closely related because of their roots in state-space modeling. It is possible to formulate state-space analogs of VAR and VEC models, as the state-space formulation is remarkably general (see Harvey (1994) for details). The estimation of state-space models sometimes relies on frequentist statistical techniques, such as maximum likelihood. Dynamic Linear Models (DLM hereafter) are Bayesian versions of state-space models. Like any other state-space model, DLM derives from the Kalman filter –not inherently a Bayesian technique but provides a method for forecasting that is consistent with the Bayesian inference (Harrison and Stevens 1976). Next we provide a brief introduction to DLMs.

Dynamic modeling has a long history, dating back to mid 1960s, in the forecasting literature (Harrison 1965). Harrison and Stevens' (1976) dynamic modeling approach comprises (i) sequential model definitions for series of observations (i.e., one model for each time period), (ii) parametric models (i.e., easy-to-interpret parameterizations model for the transition between this period's model and the next period's model), (iii) probabilistic representation of information about all parameters and observations, and hence (iv) inference and forecasting derived by posterior and predictive probability distributions. Pole, West, and Harrison (1994) provides an excellent introduction to the basic dynamic linear models with applications, whereas a full treatment of theory and methods of Bayesian time series analysis and dynamic linear models can be found in West and Harrison (1997).

In its simplest form a univariate normal dynamic linear model is defined by the following observation and state equations,

(A.1) $$Y_t = F_t\theta_t + v_t,$$

(A.2) $$\theta_t = G_t\theta_{t-1} + \omega_t,$$

where Y_t is the univariate dependent variable, θ_t is the state vector at time t, F_t is a known row vector of regressors, and $v_t \sim N(0,V)$ represents measurement and sampling errors. G_t is the state evolution matrix that defines the deterministic mapping of the state vectors between time periods t–1 and t. In most applications the state evolution matrix, G_t, is assumed to be constant over time and is set to identity matrix, which implies a rather restrictive random walk process for the state vector. In this study, we relax this assumption and infer the duration of adjustment as well as the persistent/transient nature of the series. The evolution error is distributed $\omega_t \sim N(0,W)$ and allows for stochastic deviations from the mapped values of the state vector (West and Harrison 1997). The model has a Markovian nature as the state vector varies over time following the Markov evolution equation. Sequentially arriving data points are used in the sequential updating of the summary statistics that determine the posterior distributions. Unlike standard regression, this means the estimates reflect the data that precedes the observation as opposed to the entire data sequence. Hence, the parameter estimates better relate to the state of available information at the time of the estimate. The resulting posterior distributions are used for inference about the state vector θ_t over all observations and future values of the dependent variable. Assuming normality of the initial state vector θ_0 and assuming that the only information used in updating is the set of observed values of the dependent variable (Y_t) and the independent variables (in F_t), one obtains a closed model, wherein the information is updated via $D_t = \{D_{t-1},Y_t\}$ and jointly normally distributed Y_t and θ_t. The sequential updating is based on the known Kalman equations (see West and Harrison (1997) for details).

The extension of the normal dynamic linear model to the multivariate case is straightforward: in (A.1 and A.2), Y_t, and v_t, become vectors, and F_t becomes a matrix of regressors. The normal dynamic linear model –univariate or multivariate- can further be extended by introducing deterministic terms in the evolution equation as shown in Equation (A.4).

(A.3) $$Y_t = F_t\theta_t + v_t,$$

(A.4) $$\theta_t = G_t\theta_{t-1} + h_t + \omega_t.$$

These models are known as transfer function DLMs, wherein non-stochastic sources of variation are allowed to influence the state vector. Through this state vector the new source of variation is transferred to the dependent observations. The model specified in this paper belongs to this class of DLMs.

Appendix B: Model Estimation

For a given brand j (j = 1,…,J) Equations (2)-(6) can be combined in a single model and written as

(B.1a) $$Y_t = F_{1t}\Theta_{1t} + F_{2t}\Theta_2 + \upsilon_t,$$

(B.1b) $$\Theta_{1t} = G\Theta_{1t-1} + h_t + \omega_t$$ .In Equations (B.1a) and (B.1b) Y_t is a $(3S+K)\times 1$ vector of dependent variables including log sales, log regular price, log price index, and K (= 4 in the application) marketing mix variables. F_{1t} is a $(3S+K)\times N$ matrix of regressors where N (= 7 in the application) is the number of explanatory variables with a time varying parameter (brand and store intercepts, log regular price, and advertising). Θ_{1t} is a $N\times 1$ vector of brand specific time varying parameters, υ_t is a $(3S+K)\times 1$ vector of observation equation errors. F_{2t} is a $(3S+K)\times M$ matrix of regressors with non-time varying parameters, kept in Θ_2 vector of size $M\times 1$. G (= diag([λ^{α} λ^{β} 1 $\lambda^{\xi 1}$ $\lambda^{\xi 2}$ $\lambda^{\xi 3}$ $\lambda^{\xi 4}$])) is a $N\times N$ matrix defining system evolution, and ω_t is a $N\times 1$ vector of system errors. The $N\times 1$ vector $h_t = \delta + Z'_{t-1}\gamma$ includes the lagged marketing mix and the system equation intercepts. Both Y_t and Θ_t have multivariate normal distributions, and so do the associated error terms. We assume that $\upsilon_t \sim N(0,V)$, where the variance matrix V, of size $(3S+K)\times(3S+K)$, is time invariant and full. Note that we correlate sales, regular price, promotional price and marketing mix error terms within each brand. This allows us to capture unobserved shocks that may cause endogeneity. The system errors are distributed multivariate normal, $\omega_t \sim N(0,W)$, where W is a diagonal matrix of size $N\times N$.

We place normal priors on all parameters of the Equations (B.1a) and (B.1b). The evolution equation (B.1b) error covariance matrix is assumed to be diagonal and we place an Inverse Gamma prior on their diagonal elements. As we allow for correlation between the observation equation error terms and the marketing mix equation error terms (B.1a), the associated error covariance matrix is full. Therefore we place an Inverse Wishart prior. Given these priors the estimation is carried out using DLM updating within a Gibbs sampler. Conditional on Θ_2, V, W, G, and h_t the time varying parameters (Θ_{1t}) are obtained via the forward filtering backward sampling procedure (Carter and Kohn 1994, Frühwirth-Schnatter 1994). The long-term marketing mix effects (γ) are estimated using a random walk Metropolis-

Hastings algorithm. Next, we derive the full conditional posterior distributions used in the sampling chain.

First, define $Y_t = [Y'_{1t}\ Y'_{2t}]'$ such that Y_{1t} includes log sales of the focal brand, and Y_{2t} includes the rest (log regular price, log price index, and *K* marketing mix variables). Also define $\Theta_2 = [\Theta'_{21}\ \Theta'_{22}]'$, and $F_{2t} = \text{diag}([F_{21t}\ F_{22t}])$, where Θ_{21} and Θ_{22} contain non-time varying parameters from the sales equation and remaining equations, respectively. As Y_{1t} and Y_{2t} are jointly normally distributed,

(B.2) $$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} \sim \left[\begin{matrix} F_{1t}\Theta_{1t} + F_{21t}\Theta_{21} \\ F_{22t}\Theta_{22} \end{matrix}, \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \right],$$

the conditional covariance matrix is given by $\tilde{V} = V_{11} - V_{12}V_{22}^{-1}V_{21}$, and the conditional mean vector (net off sales attributed to the variables with non-time varying parameters) is given by $\tilde{Y}_{1t} = Y_{1t} - V_{12}V_{22}^{-1}(Y_{2t} - F_{22t}\Theta_{22})\text{ - }F_{21t}\Theta_{21}$.

Assuming that the DLM is closed to external information at times $t \geq 1$ -i.e. given initial information D_0 at $t = 0$, at any future time *t* the available information set is simply $D_t = \{\tilde{Y}_{1t}, D_{t-1}\}$, and D_0 includes all values of V, W, G, and h_t and $\Theta_{10} \mid D_0 \sim N(m_0, C_0)$. Conditional on these parameters the solution is given by West and Harrison (1997). Prior at time *t* is $\Theta_{1t} \mid D_{t-1} \sim N(a_t, R_t)$, where the mean and the covariance matrix are $a_t = Gm_{t-1} + h_t$ and $R_t = GC_{t-1}G' + W$. One-step ahead forecast at time *t* is $\tilde{Y}_{1t} \mid D_{t-1} \sim N(f_t, Q_t)$, where $f_t = F_t a_t$ and $Q_t = F_t R_t F_t' + \tilde{V}$. Then the posterior distribution at time *t* is $\Theta_t \mid D_t \sim N(m_t, C_t)$, where $m_t = a_t + R_t F_t' Q_t^{-1}(\tilde{Y}_{1t} - f_t)$, and $C_t = R_t - R_t F_t' Q_t^{-1} F_t R_t$.

<u>Step 1:</u> Θ_{1t}|rest

In order to sample from the conditional distribution of Θ_{1t} for each brand *j*, we adopt the forward filtering, backward sampling algorithm proposed by Carter and Kohn (1994) and Frühwirth-Schnatter (1994). The sampling of system parameters starts with the standard DLM updating. For *t* = 1,…,*T* we apply forward filtering to obtain the moments, m_t and C_t. At *t* = *T* we sample a vector of system parameters from the distribution $N(m_t,C_t)$, then we sequence backwards for t = T-1, …, 1 sampling from $p(\Theta_{1t} \mid \Theta_{1t+1},\Phi) \sim N(q_t^*,Q_t^*)$, where $q_t^* = m_t + B_t(\Theta_{1t+1} - a_{t+1})$, $Q_t^* = C_t - B_t R_{t+1} B_t'$, and $B_t = C_t G' R_{t+1}^{-1}$. For the starting values of time varying parameters, we use m_0 = [0 -2 0 0 0 0 0], and set the initial variance C_0 to I_N.

Step 2: V|rest

For each brand j, we allow for correlations between all error terms and place an Inverse Wishart prior on the error covariance matrix. We use a diffuse prior for V that has a prior mean-diagonal element S_{V0} = $.001I_{(3S+K)}$ and set the prior degrees of freedom n_{V0} to $(3S+K)+2$. Then the full conditional posterior distribution has degrees of freedom $n_{V1} = n_{V0} + T$ with a variance matrix given by

$$S_{V1} = S_{V0} + \sum_{t=1}^{T} (Y_t - F_{1t}\Theta_{1t} - F_{2t}\Theta_2)(Y_t - F_{1t}\Theta_{1t} - F_{2t}\Theta_2)'.$$

Step 3: Θ_2|rest

In order to obtain the conditional posterior distribution of the non-time varying parameters for each brand j we define $Y^*_t = Y_t - F_{1t}\Theta_{1t}$ and $V_T = V \otimes I_T$. We place a diffuse Normal prior on the parameters, $\Theta_2 \sim N(\underline{\mu}_{\Theta 2},\underline{\Sigma}_{\Theta 2})$, where $\underline{\mu}_{\Theta 2} = 0$ and $\underline{\Sigma}_{\Theta 2} = 1000I_M$. Then the full conditional posterior is $\Theta_2 \sim N(\overline{\mu}_{\Theta 2}\overline{\Sigma}_{\Theta 2})$, where $\overline{\mu}_{\Theta 2} = \overline{\Sigma}_{\Theta 2}\{\underline{\Sigma}_{\Theta 2}^{-1}\underline{\mu}_{\Theta 2} + [F_{2t}V_T^{-1}Y_t^*]\}$, and $\overline{\Sigma}_{\Theta 2} = \{\underline{\Sigma}_{\Theta 2}^{-1} + [F_{2t}V_T^{-1}F'_{2t}]\}^{-1}$.

Step 4: W|rest

For each brand j, we assume that the system equation error covariance matrix is diagonal, and place an Inverse Gamma prior on the diagonal elements of this matrix, with $n_{W0}/2$ degrees of freedom and a scale parameter of $S_{W0}/2$. The full conditional posterior distribution is also distributed Inverse Gamma with $n_{W1} = n_{W0} + T - 1$ and $S_{W1} = S_{W0} + \sum_{t=1}^{T} (\Theta_{1t} - G\Theta_{1t-1} - h_t)^2$. We use a diffuse prior with n_{W0} = 3 and S_{W0} =.001.

Step 5: λ|rest

In this step we derive the full conditional posteriors of decay parameters for each brand j and system equation i. We place a Normal prior on all parameters, $\lambda^i_j \sim N(\underline{\mu}_{\lambda ij},\underline{\Sigma}_{\lambda ij})$, where $\underline{\mu}_{\lambda ij} = 0$ and $\underline{\Sigma}_{\lambda ij} = 1000$. We first stack the observations Θ_{1it} across time in vectors Θ_{1iT} and $\Theta_{1iT\text{-}1}$, running from $t = 2,\ldots,T$ and $t = 1,\ldots,T\text{-}1$ respectively. We also stack the corresponding components of h_t in h^i_T. Then for each i we define $y_{\Theta iT} \equiv \Theta_{1iT} - h^i_T$ and $W_{iT} = W_i \otimes I_{T\text{-}1}$. Given the normal priors, and the likelihoods, the full conditional posterior distributions are $\lambda^i_j \sim N(\overline{\mu}_{\lambda ij} \overline{\Sigma}_{\lambda ij})$, where $\overline{\mu}_{\lambda ij} = \overline{\Sigma}_{\lambda ij}\{\underline{\Sigma}^{-1}_{\lambda ij}\underline{\mu}_{\lambda ij} + [\Theta'_{1iT} W^{-1}_{iT} y_{\Theta iT}]\}$ and $\overline{\Sigma}_{\lambda ij} = \{\underline{\Sigma}^{-1}_{\lambda ij} + [\Theta'_{1iT} W^{-1}_{iT} \Theta_{1iT}]\}^{-1}$.

Step 6: δ|rest

In this step we derive the full conditional posteriors of intercepts for each brand j and system equation i ($i = 1$ for intercept, $i = 2$ for elasticity). We place a Normal prior on all parameters, $\delta^i_j \sim N(\underline{\mu}_{\delta ij}, \underline{\Sigma}_{\delta ij})$, where $\underline{\mu}_{\delta ij} = 0$ and $\underline{\Sigma}_{\delta ij} = 1000$. We stack the observations Θ_{1it} across time in vectors Θ_{1iT} and $\Theta_{1iT\text{-}1}$, running from $t = 2,\ldots,T$ and $t = 1,\ldots,T\text{-}1$, respectively. We also stack the corresponding components of h_t in h^i_T. Then for each i we define $y_{\Theta iT} \equiv \Theta_{1iT} - \lambda^i_j \Theta_{1iT\text{-}1} - h^i_T$ and $W_{iT} = W_i \otimes I_{T\text{-}1}$. Given the normal priors, and the likelihoods, the full conditional posterior distributions are $\delta^i_j \sim N(\overline{\mu}_{\delta ij} \overline{\Sigma}_{\delta ij})$, where

$\overline{\mu}_{\delta ij} = \overline{\Sigma}_{\delta ij}\{\underline{\Sigma}^{-1}_{\delta ij}\underline{\mu}_{\delta ij} + [\mathbf{1}' W^{-1}_{iT} y_{\Theta iT}]\}$ and $\overline{\Sigma}_{\delta ij} = \{\underline{\Sigma}^{-1}_{\delta ij} + [\mathbf{1}' W^{-1}_{iT} \mathbf{1}]\}^{-1}$.

Step 7: γ|rest

We use a random walk Metropolis-Hastings step within the Gibbs sampler to obtain each marketing mix coefficient in the two system equations. We generate the candidate rate draw by $\gamma^{(m)} = \gamma^{(m\text{-}1)} + z$, where (m) denotes mth iteration, and z is a random draw from $N(0, \kappa I)$. We select κ such that the acceptance rate is between 20%-50% (Chib and Greenberg 1995). The candidate draw is accepted with the probability $\alpha^* = min\{1, \alpha\}$, where

$$\alpha = \frac{L\left(\gamma^{(m)} \mid \Theta_1, W, \delta, \lambda\right) p\left(\gamma^{(m)} \mid \underline{\mu}_\gamma, \underline{\Sigma}_\gamma\right)}{L\left(\gamma^{(m-1)} \mid \Theta_1, W, \delta, \lambda\right) p\left(\gamma^{(m-1)} \mid \underline{\mu}_\gamma, \underline{\Sigma}_\gamma\right)}, \quad \text{(A.3)}$$

$L(\gamma^{(.)}|\cdot)$ is conditional likelihood of Equation (A.2), and $p(\gamma^{(.)}|\cdot)$ is the prior density evaluated at each $\gamma^{(.)}$. We set $\underline{\mu}_\gamma = 0$ and $\underline{\Sigma}_\gamma = 1000$.

Appendix References

Carter, C. and R. Kohn (1994), “On Gibbs Sampling for State Space Models,” *Biometrika*, 81 (3), 541–53.

Chib, Siddhartha and Edward Greenberg (1995), “Understanding the Metropolis-Hastings Algorithm,” *American Statistician*, 49 (4) 327-335.

Fok, Dennis, Csilla Horváth, Richard Paap, and Philip Hans B.F. Franses (2006), “A Hierarchical Bayes Error-Correction Model to Explain Dynamic Effects of Price Changes,” *Journal of Marketing Research,* 42(August), 443-462.

Frühwirth-Schnatter, Sylvia (1994), “Data Augmentation and Dynamic Linear Models,” *Journal of Time Series Analysis*, 15 (2), 183–202.

Harrison, P. Jeff (1965), “Short-term Sales Forecasting,” *Applied Statistics*, 14 (2/3), 102-139.

Harrison, P. Jeff and C.F. Stevens (1976), “Bayesian Forecasting,” *Journal of Royal Statistical Society – Series B*, 38 (3), 205-247.

Harvey, Andrew C. (1994), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.

Mela, Carl F., Sunil Gupta, and Donald R. Lehmann (1997), “The Long-term Impact of Promotion and Advertising on Consumer Brand Choice,” *Journal of Marketing Research*, 34 (May), 248–261.

Pauwels, Koen, Dominique M. Hanssens, and S. Siddarth (2002), “The Long-term Effects of Price Promotions on Category Incidence, Brand choice, and Purchase Quantity,” *Journal of Marketing Research*, 39 (November), 421–439.

Pole, Andy, Mike West and P. Jeff Harrison (1994), *Applied Bayesian Forecasting and Time Series Analysis*, New York: Chapman-Hall.

Van Heerde, Harald J., Carl F. Mela, and Puneet Manchanda (2004), “The Dynamic Effect of Innovation on Market Structure,” *Journal of Marketing Research*, 41 (May), 166-183.

West, Mike and Jeff Harrison (1997), *Bayesian Forecasting and Dynamic Models*, 2nd edn, Springer-Verlag, New York.