Brand Portfolio Promotions

Anocha Aribarg
Assistant Professor of Marketing
Stephen M. Ross School of Business
University of Michigan
701 Tappan Street
Ann Arbor, Michigan 48109-1234
Tele: (734) 763-0599
Fax: (734) 936-0279
E-mail: anocha@umich.edu

Neeraj Arora
Arthur C. Nielsen, Jr., Professor of Marketing Research
School of Business
University of Wisconsin
4191C Grainger Hall
975 University Avenue
Madison, Wisconsin 53706
Tele: (608) 262-1990
E-mail: narora@bus.wisc.edu

July, 2007
Forthcoming, *Journal of Marketing Research*

The authors thank the editor and three anonymous reviewers for their many helpful comments. The authors are listed alphabetically.

Brand Portfolio Promotions

Abstract

Large firms implement brand portfolio promotions (BPP) that promote multiple brands to targeted consumers at discrete points in time. Such programs possess unique properties that require a novel model to assess their effectiveness. The authors propose a model that captures the magnitude and shape of the BPP effect for each brand in the portfolio, accounts for diminishing returns of brand exposure, and incorporates inter-temporal effect of BPPs. The model is shown to be general enough to apply to any discrete promotion where the carry-over effect duration is unknown. Several model-based metrics that allow an objective comparison of ROI from a BPP versus other forms of promotion are presented. Results suggest that a BPP, when contrasted to feature, leads to higher sales lift per household for some of the brands. The authors develop an optimal exposure allocation procedure based on the proposed model that informs a) which assortments of brands to promote across multiple BPPs and b) the exposure level for each brand.

Large manufacturing firms such as P&G, General Mills and Unilever own a portfolio of brands across multiple product categories. In recent years, it has become common for these firms to engage in corporate-level promotion programs that encompass multiple brands and categories. Unlike traditional sales promotions such as features that are distributed to the masses and controlled by a retailer, these promotions are mailed to carefully selected consumers at a predetermined time interval and coordinated across brands that the firm owns. In this paper, we refer to such a corporate-level promotion as brand portfolio promotion (BPP). For example, General Mills mails a magazine-like promotion Que Rica Vida to over two million Hispanic consumers and Serving Up Soul to the African-American target market (Wentz 2006). Included in this multi-brand promotion is information about health, wellness and recipes. Visually the copy quality of these glossy promotion materials is similar to widely circulated magazines such as *Better Homes and Gardens*. The length of a typical BPP magazine could vary between 40-60 pages and as many as fifty brands could be mentioned in each BPP—featured as a brief mention in a recipe or an article for some brands and as a full page color ad for others (Thompson 2002). Despite its promise as an alternative promotional tool, we currently know little about the effectiveness of a BPP, how it compares to traditional forms of sales promotion, and how to enhance the return on investment of a BPP program.

With these broad research questions in mind our goals in this paper are to (i) develop a modeling framework to examine the effectiveness of BPPs (ii) characterize the nature of BPP effects on sales (iii) quantify relative effectiveness of BPP to other forms of sales promotion such as feature and (iv) provide guidelines to enhance the effectiveness of a BPP program. While the print medium is one form of a BPP seen in the marketplace, its electronic counterpart is also quite common. P&G and Kellogg have electronic mail programs under the names Home Made

Simple and Kellogg Kitchens, respectively, to promote carefully selected brands from different categories. The empirical application presented in our paper involves the print medium, although the methodology applies to electronic BPPs as well.

At first glance, a BPP appears to be similar to a feature—a vertical cooperative promotion involving distribution of printed materials—because each entails a brand related communication. However, a closer examination reveals distinct differences. First, unlike feature activities that have high coverage via an avenue such as a Sunday newspaper, BPPs are targeted at a smaller number of households. Therefore, it may be infeasible to estimate BPP effects using panel data due to insufficient sample size, especially for low penetration or infrequently purchased brands. Second, unlike a feature which communicates limited information about promoted brands, a BPP includes usage related information such as recipes. As a result, consumers likely keep BPP materials for a longer period of time than features. Because of the longer "shelf life", it is important that BPP carry-over effects on brand sales be modeled appropriately. Third, given their programmatic nature, BPPs are distributed repeatedly over time. This suggests the need to model inter-temporal effects across multiple BPPs. Fourth, because a BPP simultaneously promotes a portfolio of brands, a suitable model must accommodate heterogeneity in BPP effects across brands. Finally, because each brand experiences a different level of exposure[1], possible diminishing returns of brand exposure needs to be explicitly modeled. Interestingly, these properties point to the hybrid nature of a BPP because it resembles sales promotion in some ways and advertising in others. Measuring BPP effectiveness therefore requires a new modeling framework that can accommodate its unique properties.

[1] Coverage is a distribution measure that accounts for the number of households that receive a BPP. Exposure, in contrast, refers to the extent of visibility a brand receives in a given BPP.

Development of a modeling framework to accurately capture and understand BPP effects is an academic challenge that is also of great interest to practitioners because accurate ROI assessment is necessary for the long term success of a BPP program. It is crucial for firms to be able to gauge sales lift as a result of BPPs vis-à-vis other promotion, such as feature, that has a longer history and a demonstrated effect on ROI. Brand managers, in particular, have a great interest in accurate assessment of BPP effectiveness because they are typically asked to contribute a portion of their total promotion budget for a BPP. A suitable modeling approach should not only measure BPP effectiveness, but also inform how to best allocate limited space across a finite number of brands over multiple BPPs. Practical questions for effective management of a BPP program include a) which assortments of brands to promote and b) the extent of exposure for each brand.

To achieve our research goals, we develop a modeling framework with a system of equations with vector autoregressive error components that link a BPP program targeted at individual households to weekly store sales. In order to accurately assess effects of a BPP program, the model captures the magnitude and shape of each BPP's effect at the brand level. We propose the use of a gamma density function to characterize the shape of BPP carry-over effects because it offers distinct advantages in terms of shape flexibility, parsimony and ease of interpretation. Because of this flexibility, the model could be applied to any form of discrete promotion where the effect duration is unknown. The model accommodates different levels of exposure across brands, accounts for diminishing returns of exposure, and captures inter-temporal effect across multiple BPPs. To incorporate store-level heterogeneity, we impose a hierarchical Bayes (HB) structure to the model.

The proposed model and optimal exposure allocation procedure are tested using a unique data set provided by a multinational firm that offers brands across over fifty product categories. We show that an Erlang-2 density function best characterizes the shape of BPP carry-over effects. The results demonstrate heterogeneity in magnitude and shape of the BPP effect across brands. Evidence in support of diminishing returns for brand exposure and inter-temporal BPP effect also exists. We demonstrate the generality of our model by using it to estimate feature effects. Unlike a feature that exhibits an instantaneous effect, we find that a BPP has a non-monotonic carry-over effect after the initial launch. In addition, we find that a BPP can lead to higher sales lifts per household than a feature in some categories. Finally, we illustrate how the model could be used to optimally allocate exposure across brands and BPPs to improve the overall profitability of the BPP program.

The remainder of the paper is structured as follows. We first develop our modeling framework by relating it to prior research. Then we empirically test the proposed model and provide analyses about characteristics of BPP effects, as well as a comparison of BPP and feature effects. Next, we present an optimal exposure allocation framework and an illustration of profit gains from our suggested approach. We finally conclude with a discussion of our contributions, some limitations, and avenues for future research.

CONCEPTUAL AND MODEL DEVELOPMENT

Sales promotions are discrete activities that tend to have short-term and immediate effects on sales (Neslin 2002, P.XI first paragraph). Because of increasing concerns about undesirable effects of price-oriented sales promotion on brand loyalty and price sensitivity (Jedidi, Mela, and Gupta 1999; Mela, Gupta, and Lehmann 1997) and advertising clutter (Brown and Rothschild 1993; Keller 1991; Pieters and Bijmolt 1997), several companies have resorted to new promotion methods that are less price-oriented and better targeted (Ansari and Mela 2003; Zhang and Krishnamurthi 2004). We view a brand portfolio promotion as an example of such a targeted sales promotion. While BPPs are typically delivered via regular or electronic mails, current direct mail response models (Gönül and Shi 1998; Gönül, Kim, and Shi 2000) cannot be used to assess its impact on sales. This is because direct mail response models rely on the availability of individual customers' transaction data, which may be unavailable for BPPs. Our model therefore links store sales to BPP activities targeted at individuals.

Despite its discrete nature, we expect sales lift as a result of a BPP to last beyond a single period, a property similar to carry-over effects in advertising (Bass and Clark 1972; Clark 1976, Leeflang et al. 2000 pages 85-91; Russell 1988). The magnitude and carry-over shape of the BPP effect is also expected to vary across brands and product categories, much like what the advertising literature suggests (Assmus, Farley, and Lehmann 1984; Tellis, Chandy and Thaivanich 2000). However, because of the discrete nature of BPP activities, we directly model the shape of BPP carry-over effects using a gamma density function instead of lag terms which is an appropriate approach for capturing carry-over effects of continuous promotion activities such as advertising. Carry-over effects of a single BPP that persist after the period of promotion attempt are also different from long-term promotion effects (Jedidi, Mela, and Gupta 1999; Mela, Gupta, and Lehmann 1997) that manifest as changes in brand loyalty or price sensitivity. The programmatic nature of BPPs dictates multiple drops over time—it is therefore necessary to examine inter-temporal effect associated with BPPs.

Model Overview

We begin with an individual store level equation that links a given brand's sales to a variety of independent variables, including BPP activities. This is followed by a justification for why we use a particular specification—a gamma density function—to characterize the patterns of BPP effects across brands. Next we expand the model to include multiple BPPs. We then include important covariates driven by different levels of exposure across brands and BPPs to capture diminishing returns on exposure and inter-temporal BPP effect. Because we are interested in modeling the BPP effect on a portfolio of brands, we specify a system of equations that links BPPs to brand sales, and account for both cross-sectional and temporal correlation in sales across brands and time. Finally, the model incorporates heterogeneity in parameters across stores through the use of a Hierarchical Bayes (HB) specification.

Effect of a Single BPP

We specify demand for a given brand j at time t in a given store as follows.

(1) $$\mathrm{y}_{\mathrm{jt}} = \eta_{0\mathrm{j}} + \sum_k \eta_{kj} x_{jkt} + \tau_{\mathrm{j}} \lambda_{\mathrm{jt}} \mathrm{cov}_{\mathrm{t}} \mathrm{I}_{\mathrm{jt}} + \varepsilon_{\mathrm{jt}}$$

where y_{jt} is standardized log sales and there are k=1,…,K control variables x_{jkt} that include (i) time, a trend variable that refers to week number, (ii) a vector of dummies to capture seasonality, (iii) own price, (iv) own feature, (v) own display, (vi) competitors' price, (vii) competitors' feature and (viii) competitors' display. All independent variables involving competition are share weighted across competitors to reduce the number of parameters to be estimated (Christen et al. 1997; Kopalle, Mela, and Marsh 1999; Wittink et al. 1988). η_{kj}, k=0, 1,…,K are brand-specific parameters associated with control variables, and the term $\tau_{\mathrm{j}} \lambda_{\mathrm{jt}} \mathrm{cov}_{\mathrm{t}}\, \mathrm{I}_{\mathrm{jt}}$ captures BPP effects that we discuss in detail next. The specification of the error term ε_{jt} will be discussed later.

BPP Coverage Measure

Three pieces of information with regard to BPP circulation activities are relevant: Whether a brand was engaged in a BPP? For a give store, what fraction of the households in its trading area received the BPP? And how many households shop at that store? In Equation 1, variable I_{jt} indicates whether brand j is involved in the BPP activity in week t. Variable cov_t denotes coverage of the BPP in week t, and varies between 0 and 1. A value of 0.5 indicates that 50% of the households get the BPP. For all brands that appear in a given BPP, the cov_t variable is the same. Because different stores are associated with different number of households, we adjust cov_t by an appropriate scaling factor that incorporates the number of households who reside in the store's trading area[2]. We expect a BPP to have carry-over effects that last longer than the week in which it is launched. To capture such carry-over effects (explained in the next section), we need to allow I_{jt} and cov_t to remain the same after the BPP is launched for a longer time period than what we expect the carry-over duration to be. We use a 52-weeks time period because of the low possibility that a BPP effect would last beyond one year. This specification is different from how a mass sales promotion such as a feature gets treated in most marketing models—a feature effect is expected to last only a week and coverage is assumed to be 100%.

Shape of BPP Carry-over Effects

A BPP is likely to have a longer shelf-life in a recipient's kitchen (or inbox). It is therefore reasonable to expect that households may be induced to buy brand j during any shopping trip after receiving the BPP. This suggests the need for a flexible functional form to accommodate variation in the duration of BPP effects (i.e., carry-over effects) across brands. While the shape of the BPP effect overall may be dictated by when the target households choose to act on a BPP, the shape heterogeneity across brands may also be affected by latent factors such as inter-purchase time and category-specific variety seeking behavior. In Equation 1, the

[2] The scaling factor = #households in each store's trading area/10,000. Given that each store in our sample is associated with fewer than 10,000 households, a scale factor takes values between 0 and 1.

impact of BPP on sales is captured jointly through τ_j and λ_{jt}, which are brand specific. Parameter τ_j captures the magnitude of sales lift associated with the BPP and λ_{jt} allows the shape of this effect to vary over time. This λ_{jt} parameter can then be represented by a gamma pdf as follows.

(2) $$\lambda_{jt} = \frac{l^{\alpha_j - 1} e^{-l/\beta_j}}{\Gamma(\alpha_j)\beta_j^{\alpha_j}}, \text{ where}$$

the index l is different from the time index t; l is equal to 1 for the week in which the BPP is launched, 2 for the week after, and so on. The parameter α_j is the shape parameter and β_j is the scale parameter of the gamma pdf associated with brand j. To visualize the joint impact of τ_j and λ_{jt}, one could define $\theta_{jt} = \tau_j \lambda_{jt}$ for each brand and plot θ_{jt} over time t.

The use of gamma function has some similarities to previous research that uses transition time distributions to characterize coupon redemption (Lenk and Rao 1995) and inter-purchase time (Schmittlein and Morrison 1983) for a given consumer. Let us assume that (i) shopping trips for brand j follow a Poisson process with rate $\frac{1}{\beta_j}$, and (ii) α_j shopping trips are required before a consumer actually buys brand j. Mathematically the transition time to buy brand j then follows a gamma distribution with shape parameter α_j and scale parameter β_j. Behaviorally, such a data generating mechanism appears reasonable because a consumer could choose to buy brand j after a period of time after browsing the BPP a few times. That is, a BPP-induced brand purchase can occur during any shopping trip after receiving the BPP.

The use of the gamma density functional form in our context provides several important benefits. First, the gamma pdf allows us to capture sales lift over time parsimoniously because it requires only three parameters to be estimated: τ_j, β_j and α_j. An alternative approach, for example a dummy variable for each period, would require a lot more parameters without any precise direction for how many dummies to use. Second, the gamma pdf accommodates different possible patterns of sales lifts over time. The gamma function with α_j =1 (i.e., exponential) reflects BPP's carry-over effects that are monotonically decreasing over time. In contrast, the gamma function with α_j =2 (i.e., Erlang-2) is a lot more flexible. In addition to a monotonic shape, it also allows a non-monotonic BPP effect where the brand first experiences an increase, followed by a peak, and then a decline in sales lift. Our empirical results further demonstrate the value of an Erlang-2 distribution.

Lastly, the gamma pdf properties allow us to draw several meaningful interpretations of the parameter estimates. Given $\int_l \lambda_{jl} dl = 1$, we can interpret τ_j as total sales lift associated with brand j. The parameter λ_{jt} then captures how this lift is distributed over time. As a result, we can directly compare the magnitude of the BPP effect across brands. The higher the value of τ_j, the higher the sales lift. Because sales lifts are spread out over time, it may be of interest to also derive the mode of the lifts and the impact duration. Using the gamma pdf, the mode of the lift is $(\alpha_j - 1)\beta_j$ and 95% impact duration can be easily evaluated by calculating the inverse cumulative probability function $F^{-1}(.95 \mid \alpha_j, \beta_j)$.

Cumulative Effect of Multiple BPPs

The model so far captures the effect of a single BPP and we need to incorporate the effect of M consecutive BPP's. Therefore, we now extend Equation 1 to accommodate multiple BPPs (m=1, 2, …., M) over time. We accomplish this by summing the effect of individual BPPs. Because at a particular time t, effects of multiple BPPs may coexist, Equation 1 is rewritten as:

$$\text{(3)} \qquad y_{jt} = \eta_{0j} + \sum_k \eta_{kj} x_{jkt} + \sum_m \tau_j^m \lambda_{jt}^m \text{cov}_t^m I_{jt}^m + \varepsilon_{jt}, \text{ where } \lambda_{jt}^m = \frac{l_m^{\alpha_j - 1} e^{-l_m/\beta_j}}{\Gamma(\alpha_j)\beta_j^{\alpha_j}}$$

We allow BPP coverage (cov_t^m) to be different across BPPs, and the lift parameter τ_j^m is also allowed to vary by BPP. BPP coverage is likely to exhibit variation in the data for reasons explained in the empirical section and τ_j^m is also allowed to vary in order to account for variation in brand exposure across BPPs. We discuss this in detail next.

Accounting for Exposure Covariates

Past research (see Little 1979 for an excellent review) shows that over time advertising effects may exhibit variation because of: 1) changes in media and copy; 2) diminishing returns and 3) wear-out. Below we highlight why these factors are relevant for BPPs.

For a BPP, the exposure or overall visibility of one brand may be different from another brand. Variation in exposure is driven by the manner in which a brand appears in a BPP (e.g. a print ad, suggestion for contexts in which it could be used, an article discussing a brand). The main effect of exposure on BPP effectiveness is important to recognize because higher exposure is expected to result in higher sales (Assmus, Farley, and Lehmann 1984; Doyle and Suanders 1990; Naik, Matrala, and Sawyer 1998). At the same time, the diminishing returns argument from the advertising literature (e.g., Feinberg 1992; Little 1979) suggests that as the level of brand exposure reaches a certain level, its incremental contribution to sales should decline. Also, the wear-out argument from the advertising literature (e.g., Bass et al. 2006; Eastlack and Rao 1986; Greenberg and Suttoni 1973; Simon 1982) suggests that inter-temporal effect, caused by the presence of the same brand in consecutive BPPs, may dilute BPP effectiveness. Next, we incorporate these key covariates that are likely to moderate the BPP effect on sales.

With separate parameters to capture magnitude (τ_j^m) and shape (α_j and β_j) of sales lifts, we parameterize τ_j^m as a function of two covariates of interest: brand exposure and inter-temporal effects.

(4) $$\tau_j^m = \tau_{j0} + \gamma_1 \text{ Exposure}_j^m + \gamma_2 (\text{Exposure}_j^m)^2 + \gamma_3 \text{ Exposure}_j^{m-1}$$

A squared term of exposure is included in Equation 4 to account for diminishing returns to exposure; the parameter γ_1 is expected to be positive and γ_2 negative. To capture the inter-temporal effect, we include Exposure_j^{m-1}, the exposure level of brand j in the previous BPP; the parameter γ_3 is expected to be negative[3].

Note that equation 4 allows us to efficiently estimate exposure and inter-temporal effects by pooling information across brands and BPPs. We allow τ_{j0} to be brand specific and the $\gamma_.$'s are not specific to a brand or a BPP. The equation helps us capture the effect of a BPP (τ_j^m) at the brand level *and* at the BPP attempt level. This is accomplished without directly estimating each τ_j^m which can be infeasible when the number of BPPs (M) and brand (J) is large. Collectively, Equations 3 and 4 capture i) different exposure levels across brands; ii) diminishing returns of exposure to sales; iii) the inter-temporal effect of a previous BPP on a current BPP and iv) the effect of multiple BPPs on sales.

Full Model for a BPP Program

To complete the model, the system of brand-specific J equations is specified jointly as a seemingly unrelated regression (SUR) model to account for possible cross-sectional correlation

[3] We also tried to include Exposure_j^{m-2} but its effect is not significant.

in sales across brands (Zellner 1987). Equations 3 and 4 have both linear and non-linear parts as seen below:

$$X_j = \begin{pmatrix} x'_{j1} \\ \vdots \\ x'_{jT} \end{pmatrix} \text{ and } W_j = \begin{pmatrix} w'_{j1} \\ \vdots \\ w'_{jT} \end{pmatrix}, \text{ where}$$

(5) $$x'_{jt} = (1, Time_t, Q'_t, price_{jt}, feature_{jt}, display_{jt}, \ldots, \sum_m \lambda^m_{jt} \operatorname{cov}^m_t I^m_{jt})$$

$$w'_{jt} = (\sum_m \lambda^m_{jt} \operatorname{cov}^m_t I^m_{jt} Exposure^m_j, \sum_m \lambda^m_{jt} \operatorname{cov}^m_t I^m_{jt} (Exposure^m_j)^2, \sum_m \lambda^m_{jt} \operatorname{cov}^m_t I^m_{jt} Exposure^{m-1}_j)$$

Then, the system of aggregate demand equations is given by

(6) $$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_J \end{pmatrix} = \begin{pmatrix} X_1 & & & & W_1 \\ & X_2 & & & W_2 \\ & & \ddots & & \vdots \\ & & & X_J & W_J \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_J \\ \gamma \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_J \end{pmatrix}, \quad \varepsilon_{jt} = \phi_j \varepsilon_{jt-1} + v_{jt}$$

where Y_j, j =1, 2, ….., J is a vector of standardized log sales of brand j, μ_j is a column vector of linear parameters to be estimated which include $\eta_{.j}$ and τ_{j0}, j=1,..,J and $\gamma' = (\gamma_1, \gamma_2, \gamma_3)$. The use of standardized log sales as the dependent variable allows us to pool information across brands to estimate exposure and inter-temporal effects, and compare estimates across brands in a meaningful way.

Extending Equation 1 to capture both cross-sectional and temporal correlation in sales across brands and time, we specify the error term ε_{jt} to follow a stationary VAR(1) process (Chib and Greenberg 1995). This error specification subsumes two error components: i) the vector $v_t \sim N(\underline{0}, \Sigma)$ capturing cross-sectional correlation across brands at time t and ii) the vector ϕ_j, a component of Φ capturing serial correlation between brand j at time t and all brands at time t-1. $E(v)=0$ and $E(vv)^{'} = \Sigma \otimes I_T$, where I_T is a T×T unit matrix, and Σ is a positive definite J×J matrix. Φ is a J×J matrix with characteristic roots inside the unit circle.

The model specification up to Equation 6 does not take into account that data were collected from multiple stores s. To capture parameter heterogeneity across stores in an efficient manner, we impose a hierarchical Bayes (HB) structure on the model. Given little within-store variation in coverage and a large number of stores with low coverage, we only account for the store heterogeneity in control variables. Referring back to Equation 1, the vector $\eta_s^{'} = (\eta_{10},..,\eta_{1K},.....,\eta_{J0},..,\eta_{JK})$ is allowed to vary by store and $\eta_s \sim MVN(\overline{\eta}, D_\eta)$. We select this distribution of heterogeneity because it is commonly used in marketing (Rossi and Allenby 2003). While other functional forms could have been used, prior work by Andrews, Ainslie and Currim (2002) and Ansari and Mela (2003) suggests that the multivariate normal is a good approximation to different shapes of heterogeneity distributions (e.g., gamma and bi-modal distributions). Specifying the linear part of the model as in Equation 6 helps us simplify the estimation procedure using MCMC methods. Please see the web appendix for detailed estimation procedure.

EMPIRICAL FINDINGS

Data Overview

Data for this research were provided by a multinational company that offers brands across more than fifty categories[4] and ACNielsen. Most of these brands are present in traditional channels such as grocery stores, drug stores and mass merchandisers. In a corporate-level effort to promote directly to consumers, the company periodically distributes BPPs to over a million households in a western country. This BPP program had been in place for several years at the

[4] In order to protect its identity, we cannot reveal the company name as well as brand names.

time of our analysis. Each BPP contains information about the company's products such as suggestions for how the products could be used and brand-specific ads.

We obtained weekly store-level data for ten brands over a 210-week period. The ten brands are from different product categories, routinely included in the BPP and selected randomly. The data include information on brand sales, prices, feature, display, *and* BPP coverage associated with each store. Data on sales, prices, feature and display for competitive brands that accounted for 80% of the category sales for all ten categories were also available. BPP coverage was calculated by mapping each store trading area to addresses of households around it. It is equal to the total of number of households receiving the BPP divided by the total number of households in the store trading area.

Because of their targeted nature, the BPP materials were sent to households more likely to consume in the categories in which the firm has presence. The targeting was done at the zip+4 level. Existing data were first used to uncover variables (e.g. demographics) most likely to explain consumption. Based on a scoring algorithm, zip+4 areas that scored high on these variables were included in the sample. All model-based conclusions therefore apply to a sample selected using this targeting rule. The stores in our data are spread over ten different regions of the country and only those stores (n=126) that carried all ten brands during the 210-week period were included in the study. Fourteen BPPs were distributed successively during the 210-week period with the periods between two BPPs ranging from 9-14 weeks. The first 56 weeks is the period before the BPP program was launched.

Summary Statistics

Table 1 presents summary statistics for and characteristics of the data we use for model calibration. Some of the brands have a dominant share (A, B, C and D) and others do not (G, H, I, J). Large variation in the variables reported is observed. The average sales across brands varied from 38.8 to 277.3 pounds/week. Large variation in price (per pound), feature, and display is also observed. Brand C was featured the most (24%) and Brand A the least (6%). Brand I was displayed a lot (25%) whereas Brand J was rarely on display (3%). Private label brands are strong competitors for certain brands (E, F, J) and number of competitors vary substantially. Seven of the ten brands belong to categories that could be labeled as ingredients, the rest do not.

Insert Table 1 Here

The average coverage (i.e. percent of households in the store's trading area that received BPP) across the fourteen BPPs and ten regions was 9% with a standard deviation of 10%, and a range of 0 to 88%. Large variation in coverage is observed across the 14 BPPs. Region 4, for example, had high coverage of over 30% for BPP5 and low coverage of less than 10% for several BPPs. Large variation in coverage across regions and multiple campaigns is desirable in order to estimate the model parameters (τ and β) that inform BPP effectiveness.

Each brand could appear in a BPP in a variety of ways, such as mentions in articles, suggestions for usage, pictures in the table of contents, full-page or half-page ads, etc. The company developed a point system to measure the exposure level of a brand by assigning a fixed number of points to these different ways a brand can possibly appear in a BPP. More points are associated with higher exposure. For instance, a back cover ad received more points than a ¼ page ad inside the BPP. Similarly, a brand-specific ingredient shown in a recipe was given more points than a minor mention in an article that merely promoted the category. Given that a brand could appear multiple times, and in different ways in a BPP, the measure for brand exposure that ranged between 0 and 87 was calculated by summing points associated with each brand. The average exposure across fourteen BPPs and ten brands was 17 with a standard deviation of 16.

On average across BPPs, the least exposed brands were C (6.3) and I (7.4) and the most exposed brands were G (42.8) and J (32.2). In addition to inter-brand variation, there was substantial variation in exposure levels of a brand across BPPs. Across the fourteen BPPs, the exposure range for Brand G was 19.1 to 83, and for brand I was zero to 28.8. Such large variation in exposure across brands and BPPs is desirable in order to estimate the γ parameters (Equation 4) that assess the link between exposure and BPP effectiveness.

Model Estimation: Preliminary Tests

We began by fitting a simple aggregate model that attempts to detect BPP effects at four arbitrary points in time (1 week, 1 month, 2 months and 7 months) after a BPP drop. The results showed that both the magnitude and shape of the BPP effect vary by brand. Our proposed model specification captures this heterogeneity in BPP effects across brands using a parsimonious specification that does not require a parameter for each week. In testing our model, we learned that estimating α and β at the same time could be difficult. The reason is that different combinations of α and β values can give rise to very similar patterns of lifts (e.g., α =2 and β =0.5 can give rise to a similar pattern when α =1 and β =0.5). Perhaps because of this reason, previous research (e.g. Gupta 1988, 1991) that uses the family of gamma distributions to characterize transition times has limited its focus to exponential (α =1) and Erlang-2 (i.e., α =2) distributions. Unlike the exponential distribution which is monotonically decreasing, the Erlang-2 distribution provides greater flexibility by allowing for both monotonic and non-monotonic shapes. Preliminary empirical results (we assume ϕ_j=0 in Equation 6 in preliminary tests to save on computation time) support this greater flexibility argument. We find that the model with α =2 or Erlang-2 function (DIC[5] = -112628) provide a better fit than the model with α =1 or exponential function (DIC = -112708). Because of its ability to accommodate non-monotonic shapes, Erlang-2 captures BPP carry-over effects well.

Additional model fit comparisons also point to the suitability of an Erlang-2 specification. In these comparisons, we considered three benchmark models in which a uniform BPP effect is assumed to last: i) 1 week, ii) 8 weeks or iii) 16-weeks. The model comparison results show that our proposed model (DIC = -112628) fits better than the 1-week (DIC = -112841), 8-weeks (DIC = -112875), and 16-weeks (DIC = -112722) BPP effect models. These results suggest that the Erlang-2 specification is more appropriate than a uniform BPP effect, and provide empirical support for the flexibility advantage of the Erlang-2 specification.

To investigate the generalizability of the Erlang-2 approach for *any* form of discrete promotion, we estimate a model where we impose the Erlang-2 shape on both BPP and feature effects. To see the results visually we plot feature and BPP effect for three brands (B, D and F) in Figure 1. These results show that BPP effects first increase, reach a peak, and then taper off. In contrast, feature effects peak in the first week and dissipate quickly thereafter. The β parameters (i.e. mode of the lift distribution) for the feature effect range between 0.70 and 1.28 for the ten brands. This result is consistent with the monotonic effect seen visually in Figure 1. Nonetheless, the fit statistics show that a model with 1-week dummy (DIC = -112628) to capture feature effects performs better than that with the Erlang-2 specification (DIC = -112752). While the above findings demonstrate generalizability of the proposed model in capturing the carry-over

[5] We use deviance (Gelman et.al. 2004), D(y|θ) = -2logp(y|θ), as the measure of model fit. Because of the connection between the deviance and the Kullback-Leibler information measure, the lowest expected deviance will have the highest posterior probability for large sample sizes. We compute mean deviance based on simulated draws of the posterior distribution ($\hat{D}_{avg}(y) = \frac{1}{L}\sum_{l} D(y,\theta^{l})$).

effect pattern for any form of discrete promotion, the fit statistics show that feature effects are best modeled as a 1-week effect, which is the specification we use in our final model estimation reported next.

Insert Figure 1 Here

Parameter Estimates

After the above preliminary model tests, we obtain parameter estimates for the full model (Equation 6). The parameter estimates are obtained using the MCMC estimation algorithm described in the Web Appendix. Generating 5,000 draws, we kept every 25^{th} draw of the last 2,500 draws to compute posterior means of the parameters (Table 2). We will first briefly discuss the estimates involving control parameters such as own and competitive price, feature and display. This is followed by a discussion of model parameters involving BPP coverage and exposure.

Insert Table 2 Here

Across all brands, own price is found to be a significant predictor of sales. In particular, Brands A and G are the most elastic and Brand I the least. Across all brands, own feature and display estimates are also significant and their magnitudes vary. While competitive price effects are strong and significant, the competitive feature and display effects are mostly insignificant and weak. We also observe unexpected signs for the competitive feature effects for Brand B and Brand J. One possible explanation may be that because these two specific brands compete primarily with private labels and other small brands, competitive features had a small spill-over effect on these brands. Seasonal patterns appear to exist across all brands. Eight of the ten brands show a downward trend in sales over time. This, in part, may be driven by the fact that an increasing number of people were shopping at Wal-Mart and A.C. Nielsen is unable to measure sales for Wal-Mart stores. Finally, we observe that the diagonal values of Φ are positive and significant, indicating the presence of autocorrelation in the sales data.

All ten lift parameters (τ_o) are significantly higher than zero, and exhibit a large variation across brands. Brand D (τ_o = 10.25) exhibits the highest lift and brand A (τ_o = 2.63) the lowest. This variation in lift by brand could potentially be explained by category specific factors such as purchase cycle time, brand share, number of the competitors and private label share. For example, are frequently purchased brands likely to exhibit higher lifts because of salience effects of a BPP? Does high private label presence suggest a lower lift because of possible substitution? While the limited number of brands (n=10) included in this paper can not provide conclusive evidence, own market share (correlation=.48 p=.18) and private label share (correlation=-.4, p=.28), are good candidates for further investigation. When expanded to a larger number of brands, such analyses could uncover factors most likely to impact BPP effectiveness.

Estimates for the shape parameter (β), which is the mode of the distribution, also varies across the ten brands. Notice that brand B reaches its peak effect the fastest (smallest β of 1.93) and brand H the slowest (largest β of 5.53). To see the joint impact of these two parameters we define $\theta_t = \tau_o \lambda_t$ for each brand and plot θ_t over a 30-week period for all ten brands in Figure 2. We can see that the lift for Brand B occurs pretty early and dissipates quickly whereas the lifts for Brands F, G and H linger over a long period of time. The 95% duration impact exhibits a large range of 9 to 26 weeks. This pattern of results is consistent with our assertion that a BPP is likely to have a shelf-life beyond one week.

Insert Figure 2 Here

In the model, we also include terms to capture the effect of brand exposure (Equation 4). The results show that γ_1 and γ_2 estimates (11.00 and -2.04, respectively) are significant (probability<0.05). The negative sign for γ_2 supports the diminishing returns argument. The parameter estimate γ_3 (-5.91) is found to be significant (probability<0.05), providing evidence in support of a negative inter-temporal effect. This implies that if a brand received exposure in the previous BPP, its lift parameter value in the current BPP is lower, as compared to a situation in which the brand did not receive any exposure in the previous BPP. We attribute this effect to wear-out, although stockpiling (van Heerde, Wittink, Leeflang 2004) could also potentially manifest itself as a negative inter-temporal effect.

Insert Figure 3 Here

To see how lift and shape vary by brand and BPP, we plot the percent sales gain over the entire duration of the data in Figure 3. Across brands, variation in percent sales lifts is not only driven by differences in parameter estimates, but also by changes in BPP coverage over time. Despite its complex pattern, we can still see that the peak of sales lift for Brand B occurred earlier than those of other brands, consistent with the pattern we observed in Figure 2.

To ensure stability of the reported parameter estimates (Table 2) we conduct a split-half validation. We divide the data into two parts by randomly assigning stores to two separate groups. We estimate our proposed model for each group and compare the results with what we report in Table 2. Of particular interest are the estimates pertaining to BPP. The results show that the estimates for both the first and second halves are statistically the same (cut-off probability or significance level of 0.05) as the estimates reported in Table 2. The split-half results therefore provide evidence in support of the stability of our results.

Comparison of BPP and Feature Effects

A good model should help make decisions that enhance a firm's return on investment. Next we demonstrate how firms can use our model to derive a variety of metrics to measure BPP effectiveness. These metrics are beneficial because they provide an objective guideline for promotion allocation decisions across different types of promotion. To obtain appropriate measures of effectiveness we use store specific parameter estimates to calculate (i) baseline sales without BPP or feature (ii) sales in the presence of BPP activities only, and (iii) sales in the presence of feature activities only. Based on these calculations we create several metrics that are reported in Table 3.

Insert Table 3 Here

The results in the top half of the table pertain to BPP effectiveness. For the 126 stores in the data set, the average number of households per store is 3554. The average number of households receiving a BPP is less than 10% and varies between 299 and 320 per store. The difference between (ii) and (i) in the previous paragraph provides a direct measure of unit sales lift attributable to the BPP activities that occurred during the 152 weeks of the data. In addition to reporting the unit lift, we also report the contribution margin and percent lift per BPP. Finally, for ease of comparison, the incremental unit sales lift is also translated into unit lift per BPP/store and per BPP/household[6].

The bottom half of the table corresponds to feature effectiveness. Unlike a BPP, where households are selected based on pre-determined criteria, a feature has much wider coverage (e.g., through mail, retail stores, Sunday newspaper). For the purpose of our calculation, we assume 100% feature coverage—the average number of households receiving feature is therefore assumed to be the same as the average number of households/store. Across brands, the average number of features during the 152 weeks of data is higher than the number of BPPs. For all brands, features result in much higher overall unit lifts than BPPs. These findings are expected because of i) the higher average number of features/store in the data, as compared to the number

[6] We thank a reviewer for this idea.

of BPPs (e.g., 27 features on average versus 14 BPPs for brand G) and ii) higher coverage of feature. While the measure of unit lift/store/promotion attempt (i.e., feature or BPP) reported in Table 3 accounts for the first reason, it does not account for the second. A more meaningful measure is unit lift/household/promotion attempt. The comparison of feature and BPP based on this measure indicates that a BPP is more effective than a feature for brands C, D, E and G. The opposite is true for brands A and I. Of course, these conclusions should not be extrapolated to the general population given the sample selection process for a BPP. These results should be viewed as the comparison between the effectiveness of a promotion targeted to the masses (features) vs. that of a promotion targeted to a smaller subset of carefully selected consumers (BPPs).

Table 3 illustrates how the proposed model can be used to assess relative effectiveness of a BPP program, as compared to another promotion form such as feature. Such an assessment has direct implications for resource allocation of promotion dollars. In addition, the cumulative contribution margin for all brands (not just the ten that we studied) that appear across multiple BPPs, viewed in conjunction with the cost of producing and distributing the BPPs, could provide a direct measure of ROI for the BPP program. A substantial variation in contribution margin across brands is observed in Table 3. The model's ability to capture this variation in profitability can provide some guidance to effectively manage a BPP program. We demonstrate this important use of the model in the next section.

Optimal Exposure Allocation for a BPP Program

Several papers in marketing have investigated ways in which return on promotional dollars could be maximized (e.g. Mahajan and Muller 1986; Doyle and Saunders 1990; Feinberg 1992, 2001; Bronnenberg 1998; Naik, Raman and Winer 2005; Bass, Bruce, Majumdar and Murthi 2006). Effective management of a BPP program could also be viewed as an optimization problem where key questions include: Which brands to include in each BPP attempt? What is the extent of exposure that each brand should receive? Next we illustrate how the proposed modeling framework could be used to optimally allocate the constrained BPP space across different brands and BPPs.

Our goal in the following illustration is to develop a promotion exposure calendar for a portfolio of brands. We restrict our attention to the ten brands included in this paper and BPP5, BPP6 and BPP7. We calculate the profitability[7] of the BPP exposure calendar actually used by the company and assess ways in which it could be improved. The objective function is to maximize incremental profit for the ten brands across the three BPPs. We use a greedy search procedure (Krieger and Green 1985) because brute force search for optimal exposure from all possible combinations of brands and BPPs was infeasible.

Several differences between the actual and optimal allocation emerge (detailed results are reported in the Web Appendix). First, the optimal allocation would have resulted in an increased profit of 5.32%. Second, across the three BPPs, brands D and J should have received a lower exposure and the resulting available space could have been profitably allocated to brand A. Finally and most interestingly, Brand F, G and H should have received a much higher exposure for BPP5 and BPP7 than the actual, and should not have been included at all in BPP6. In contrast to the current exposure allocation, these results point to exposure pulsing as a more profitable BPP management strategy for some brands.

The above illustration demonstrates the value of the proposed model in assessing which brands to include in the BPPs and the relative exposure of those brands. Results suggest that opportunities to allocate BPP exposure more efficiently exist. The pulsing pattern in our

---

[7] Technical details are available from the authors upon request.

optimization exercise is primarily driven by the negative inter-temporal effect present between discrete BPP drops. This argument is more in line with Park and Hahn (1991), Naik, Mantrala, and Sawyer (1998) and Simon (1982) than the S-shaped response function argument provided by others (Feinberg 1992; Mahajan and Muller 1986; Sasieni 1971, 1989). Because of the non-linear objective function (e.g. margins, lift magnitude/shape and exposure level vary by brand), a priori prediction for the optimal exposure allocation is not possible. This suggests that much like a direct mail marketing applications, model based optimization should be field tested on a small scale (e.g. in a given region) before adopting it for larger BPP campaigns.

DISCUSSION

At a tactical level, a BPP is a promotion tool that is more targeted and provides a lift that lasts longer than traditional sales promotions. Theoretically, its effect could be attributed to increased brand awareness, knowledge of new usage contexts or greater brand liking as a result the message content. At a strategic level, a BPP could be used to exploit complimentarity between brands that a firm offers, build individual brand equity, and even fortify a firm's corporate brand. Because large CPG firms often have presence in scores of product categories, complementarity is natural (e.g., waffles and syrup, cake mix and frosting). By carefully selecting brands to include in a BPP, a firm can exploit synergies between brand sales across categories. In an effort to build an individual brand (e.g., Hamburger Helper), a brand manager can target BPPs to market segments most likely to act as opinion leaders. At a broader level, corporations that rely heavily on their corporate brand names (e.g., Kellogg) could use BPPs as a means to communicate the collective power of the brands in a single medium to the more profitable segments in the market.

For the above tactical and strategic reasons, BPPs have become increasingly popular in the marketplace. Several unique properties associated with BPPs suggest that there is a need to develop a new modeling framework to test BPP effectiveness. In this paper, we develop one such model. The model structure is parsimonious yet flexible. It successfully captures heterogeneity in the magnitude and shape of sales lift across brands as a result of BPPs, and accounts for diminishing returns and inter-temporal effect of BPP exposure. The proposed model could be applied to any form of discrete promotion where the effect duration is unknown. Several ROI-related metrics are distilled out of the model to assess relative effectiveness of a BPP against traditional forms of sales promotion. Building an optimal exposure allocation framework based on our model, we demonstrate that a firm can potentially improve profitability of its BBP program by optimally allocating exposure across brands and BPPs.

Using data involving ten brands over 210 weeks and fourteen BPPs to test the proposed model, we uncover several interesting results. First, although both features and BPPs are discrete promotions, our results suggest that unlike features, BPPs exhibit non-monotonic carry-over effects beyond the period of its launch. Our results also suggest substantial heterogeneity in the magnitude and shape of sales lifts across brands. With regard to effectiveness, we find BPPs leading to higher lifts/household/promotion attempt than features for some brands. Finally, given the discrete nature of BPP and its proneness to repetition wear-out (i.e., negative inter-temporal effect), our optimal exposure allocation exercise suggests that a pulsing exposure strategy across multiple BPPs may have been more effective for some brands.

A unique aspect of a BPP is that while it is targeted at the household level, its effect needs to be measured at the market level. An appealing aspect of such a “micro-macro” approach to measurement, modeling, and subsequent optimization, is its generalizability to some retailing practices. Retail stores (e.g., Home Depot) often send direct mail promotions—similar in a lot of ways to a BPP— to households to increase store traffic and encourage after-sales services. Our proposed model can help assess effectiveness of these programs by linking promotions to store sales. Another interesting context involves retailers that have presence in multiple channels (e.g. Lands' End products can be ordered online, at a physical store, and via a catalog). A firm can use our model to gauge how promotion in a catalog may spill-over to physical stores. In general, our model applies to contexts in which individual consumer response to discrete promotion activities may not be observable, but promotion coverage and sales data at a more aggregate level (e.g. store) are.

There are several limitations and opportunities for future research. While the scaling factor in Equation 1 accounts for sales lift differences because of store size (i.e. for a given lift parameter τ, larger stores experience a larger sales lift), it is plausible that the parameter τ is also store specific. In our analyses, we were unable to estimate store-specific lift parameters because of a large number of stores with low coverage (33% of the stores have less than 5% coverage) and little within-store variation in coverage among those stores. Given our data limitation, future research may benefit from conducting a field experiment where the variation in coverage across and within stores can be manipulated such that store-level life parameters can be estimated. Heterogeneity in lift parameters could be easily incorporated by treating the lift parameter in a manner similar to the control variables (see Step 1, Web Appendix).

As suggested earlier in the paper, the model structure allows for further predictive analyses involving lift magnitude and duration by expressing τ =f(covariates) and β =f(covariates). Relevant lift magnitude covariates permit an investigation of brand-specific questions such as: Is the BPP effect moderated by presence of private labels? Is BPP more effective in mature or new categories? Lift duration covariates also allow us to explore questions such as: Do brands with shorter inter-purchase times tend to spike sooner? When accompanied by a coupon, does a BPP effect occur faster? While we do not conduct such brand-level analyses because of the small sample size (n=10 in our data), the model structure permits easy investigation of these questions given data on all brands that were included in a BPP.

We found that there is very little variation in times between BPPs (range = 8-14 weeks; mean = 11; std = 2) in our data. As a result, we could not incorporate the potential moderating impact of inter-BPP time on BPP effect. For example, it is certainly plausible that the effect of a current BPP may be mitigated if it is launched too close in time to the previous BPP. We recognize inter-BPP time as an important decision variable in managing a BPP program, and encourage future research to study its effect on overall program profitability. Such a model could also lend itself nicely to a dynamic optimization framework (Mesak and Zhang 2001) that could help determine the optimal inter-BPP time.

The optimal exposure allocation for a BPP program deserves further investigation. While the profit gain of 5% suggested by our exercise is modest, refinements to the proposed approach may result in greater gains. For example, while the ten categories used in our analyses were independent, inter-related categories that are likely to exhibit complementarity are also present in a BPP. For such categories, brand portfolio level synergies could be captured by including cross BPP elasticity terms in Equation 1. Spill-over effects of the corporate brand name, or the family brand names, that are quite prominent in a BPP may also be accounted. On the flip side, many companies operate with multiple brands within a category, and substitution effects may have important implications for their BPP programs. Careful modeling of these possible effects is likely to suggest ways in which a BPP program could be managed more profitably.

Finally, in this paper we do not model the process by which households are selected. With additional data related to household selection, a structural model to assess BPP effects can be developed. According to the Lucas critique, such a structural model is preferred because it better handles a potential endogeneity problem (Franses 2005). Another direction to develop a structural model is by modeling the process that dictates choice and exposure of brands that are included in a BPP.

Table 1
Summary Statistics

	Market share	Average Weekly Sales (in Pounds)	Average Price (per Pound)	Average Feature	Average Display	Private label share	Number of Competitors	Purchase cycle time	Ingredient
Brand A	0.76	130.2	0.17	0.06	0.15	0.12	3	71	Yes
Brand B	0.73	38.8	0.31	0.08	0.18	0.21	4	69	Yes
Brand C	0.7	226.6	0.49	0.24	0.17	0	3	56	No
Brand D	0.67	52.3	0.2	0.09	0.09	0.14	3	52	Yes
Brand E	0.46	137.2	0.18	0.12	0.12	0.35	3	66	Yes
Brand F	0.41	118	0.22	0.15	0.09	0.45	4	50	Yes
Brand G	0.27	277.3	0.18	0.17	0.19	0.12	7	61	Yes
Brand H	0.25	68.3	0.28	0.09	0.09	0.26	16	59	No
Brand I	0.24	194.9	0.27	0.2	0.25	0.12	8	58	No
Brand J	0.22	71	0.29	0.15	0.03	0.52	3	53	Yes

Table 2
Brand Specific Parameter Estimates[1]

	Brand A	Brand B	Brand C	Brand D	Brand E	Brand F	Brand G	Brand H	Brand I	Brand J
CP related parameters										
Lift parameter (τ_0)	**2.63**	**7.68**	**5.25**	**10.25**	**5.36**	**5.09**	**2.82**	**4.51**	**4.58**	**3.44**
Shape parameter (β)	**3.53**	**1.93**	**4.15**	**3.81**	**4.68**	**4.87**	**4.84**	**5.53**	**4.23**	**4.55**
95% impact duration	17	9	20	18	22	23	23	26	20	22
Non-CP related parameters										
Trend (α_1)	**-0.06**	**-0.08**	**-0.03**	**-0.11**	**-0.06**	**-0.02**	**-0.10**	**0.06**	**-0.08**	**0.05**
Seasonality (α_2) - Q2	**-0.25**	**-0.24**	**-0.14**	**0.04**	**0.40**	0.02	**0.11**	**0.05**	-0.02	**-0.04**
- Q3	**-0.23**	**-0.20**	**-0.20**	**-0.15**	**0.19**	**0.06**	0.01	**0.10**	**-0.05**	-0.01
- Q4	**0.13**	**0.39**	**-0.09**	**0.11**	**0.38**	**0.40**	**-0.18**	**0.28**	**0.04**	**0.11**
Price (α_3)	**-10.73**	**-6.22**	**-5.46**	**-6.87**	**-7.50**	**-6.89**	**-11.89**	**-4.30**	**-3.04**	**-6.62**
Feature (α_4)	**0.36**	**0.41**	**0.31**	**0.50**	**0.59**	**0.37**	**0.32**	**0.47**	**0.56**	**0.28**
Display (α_5)	**0.22**	**0.29**	**0.37**	**0.25**	**0.31**	**0.22**	**0.22**	**0.46**	**0.36**	**0.37**
Competitive price (α_6)	**1.63**	**0.63**	-0.06	**2.14**	**1.75**	**1.43**	**3.35**	**0.37**	**1.74**	**1.49**
Competitive feature (α_7)	0.03	**0.19**	**-0.06**	0.01	0.01	0.00	0.02	**-0.03**	**-0.05**	**0.039**
Competitive display (α_8)	-0.02	0.01	**-0.06**	0.04	-0.03	-0.03	-0.02	-0.00	**-0.07**	-0.02
Diag (Φ)	**0.43**	**0.38**	**0.37**	**0.48**	**0.53**	**0.62**	**0.33**	**0.40**	**0.34**	**0.54**

[1]Estimates that are statistically significant (probability<0.05) are in **bold.**

Table 3
Comparison of BPP and Feature Effectiveness
 # Stores = 126; # Periods = 152 weeks; Average #hh/store = 3554

		Brand A	**Brand B**	**Brand C**	**Brand D**	**Brand E**	**Brand F**	**Brand G**	**Brand H**	**Brand I**	**Brand J**
BPP	Average # hh receiving each BPP	305	306	320	309	309	307	309	309	299	309
	# BPPs	10	11	11	14	14	13	14	14	11	14
	Unit lift	8629	8779	25811	12824	21121	16987	39401	9152	20572	11308
	Contribution margin ($)	7594	11677	40782	12567	20699	14439	38219	10525	28184	8707
	Percent lift/BPP[1]	0.33	1.9	0.5	0.97	0.51	0.48	0.42	0.33	0.62	0.46
	Unit lift/store/BPP	6.1	6.12	19.2	6.18	12.36	9.21	21.63	6.18	14.95	6.18
	Unit lift/hh/BPP	0.02	0.02	0.06	0.02	0.04	0.03	0.07	0.02	0.05	0.02
Feature	Average # hh receiving each Feature[2]	3554	3554	3554	3554	3554	3554	3554	3554	3554	3554
	Average # feature	9	13	39	14	18	23	27	14	33	24
	Unit lift	96511	71993	396263	63772	215677	210288	394660	91084	804263	126038
	Contribution margin ($)	84930	95751	626096	62497	211364	178745	382820	104747	1101841	97049
	Percent lift/feature[3]	32.09	47.31	25.88	41.46	41.36	39.73	23.36	26.19	87.79	27.25
	Unit lift/store/feature	106.62	71.08	106.62	35.54	106.62	106.62	142.16	71.08	248.78	71.08
	Unit lift/hh/feature	0.03	0.02	0.03	0.01	0.03	0.03	0.04	0.02	0.07	0.02

[1]Percent lift per BPP = [(aggregate units lift for all BPPs/baseline sales (no feature or BPP))*100)]/#BPPs

[2]Assumes feature coverage=100%

[3]Percent lift per feature = (aggregate units lift in weeks with features/baseline sales (no feature or BPP) in those weeks)*100

Figure 1
Shape Comparison of Feature and BPP Effects

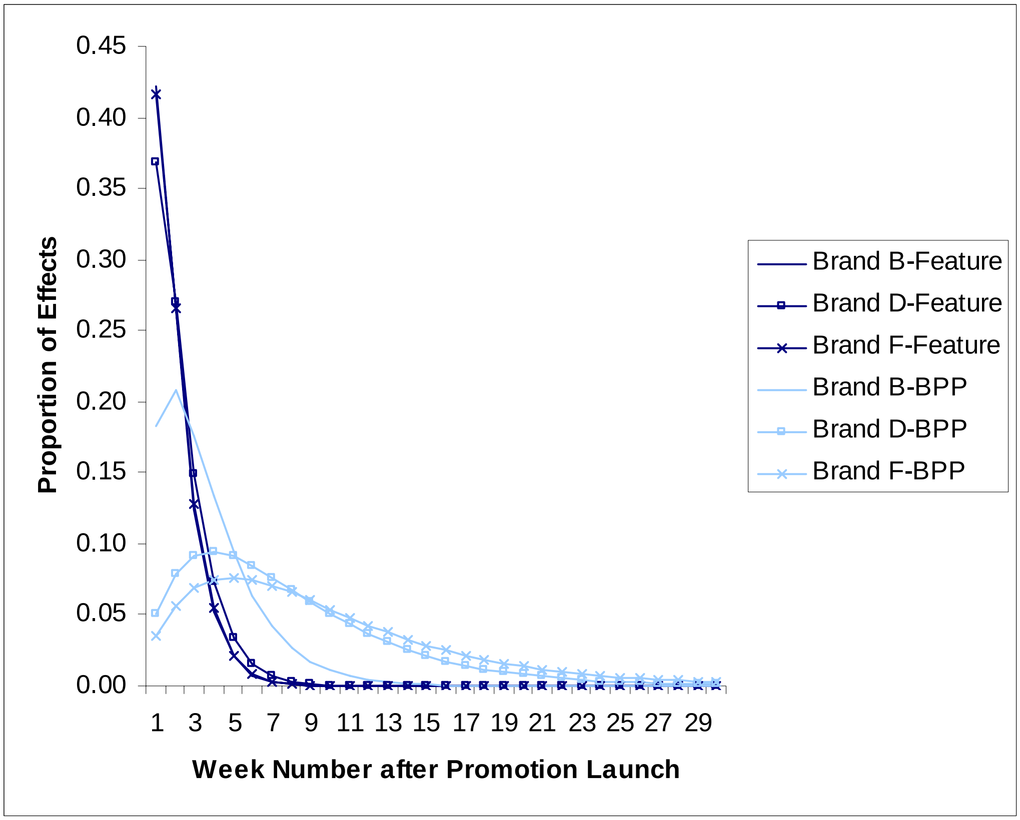


Figure 2
Lift and Shape Variation by Brand

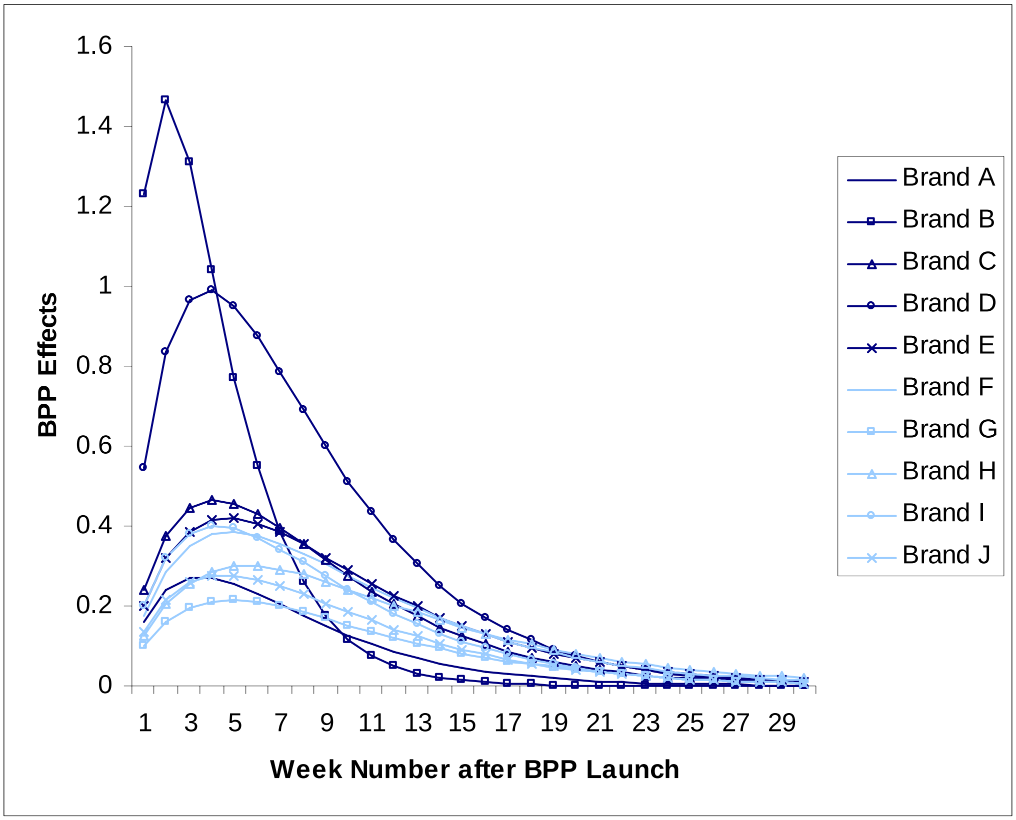
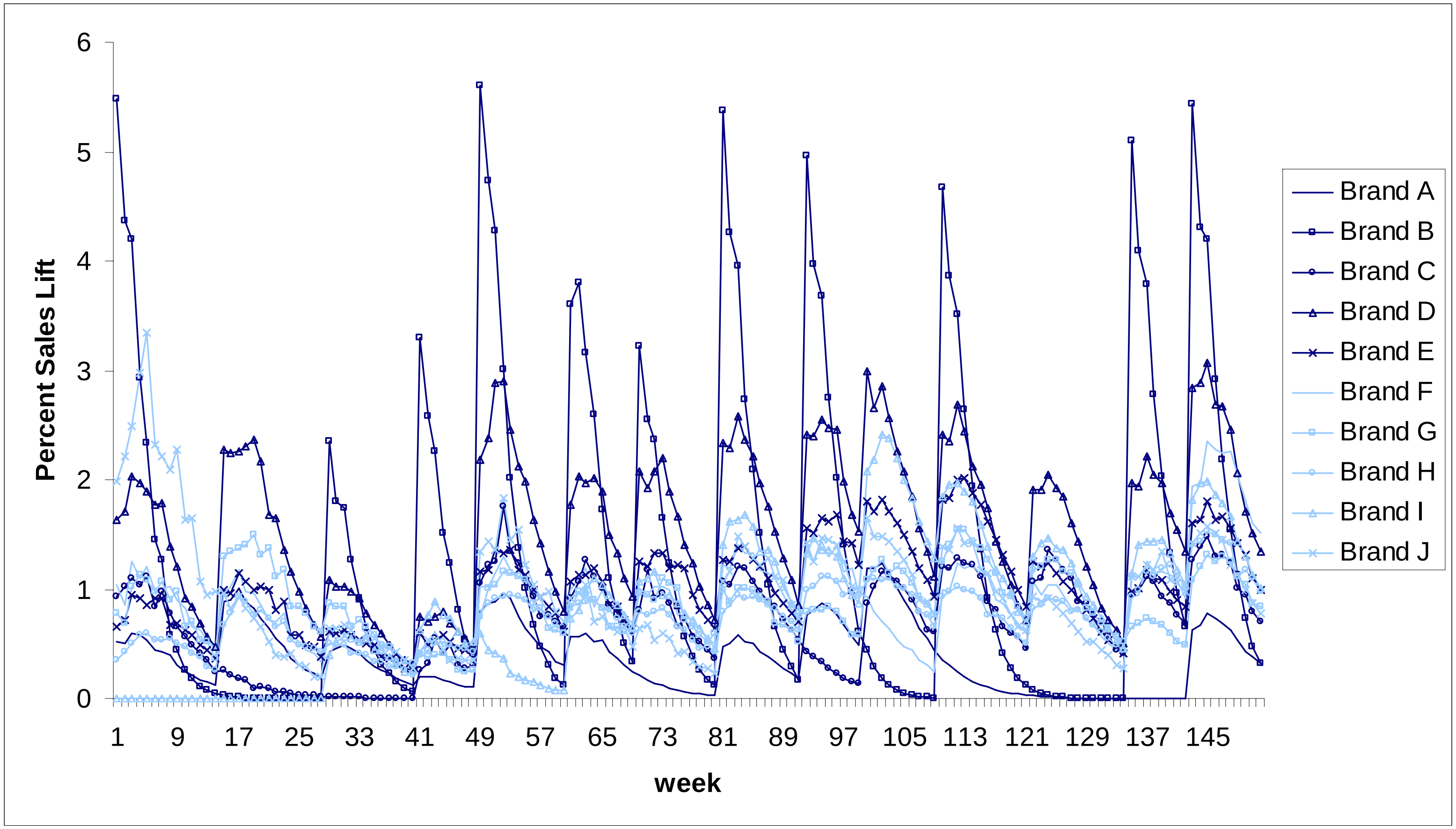


Figure 3
Lift and Shape Variation by Brand and BPP



References

Ansasi, Asim and Carl F. Mela (2003), "E-Customization," *Journal of Marketing Research*, 40 (May), 131-45.

Andrews, Rick L., Andrew Ainslie, Imran S. Currim (2002), "An Empirical Comparison of Logit Choice Models with Discrete versus Continuous Representations of Heterogeneity," *Journal of Marketing Research,* 39 (November), 479-488.

Assmus, Gert, John Farley, and Donald Lehmann (1984), "How Advertising Affects Sales: Meta –analysis of Econometric Results," *Journal of Marketing Research*, 21 (February), 65-74.

Bass, Frank M., Norris Bruce, Sumit Majumdar, and B.P.S. Murthi (2006), "A Dynamic Bayesian Model of Advertising Copy Effectiveness in the Telecommunications Sector," *Marketing Science*, forthcoming.

Bass, Frank M. and D.G. Clarke (1972), "Testing Distributed Lag Models of Advertising Effects," *Journal of Marketing Research*, 9 (August), 298-308.

Bronnenberg, Bart J. (1998), "Advertising Frequency Decisions in a Discrete Markov Process under a Budget Constraint," *Journal of Marketing Research*, 35 (August), 399-406.

Brown, Tom J. and Michael Rothschild (1993), "Reassessing the Impact of Television Advertising Clutter," *Journal of Consumer Research*, 20 (June), 138-46.

Chib, Siddhartha and Edward Greenberg (1995), "Hierarchical Analysis of SUR Models with Extensions to Correlated Serial Errors and Time-varying Parameter Models," *Journal of Econometrics*, 68, 339-60.

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

Clarke, D.G. (1976), "Econometric Measurement of the Duration of Advertising Effect on Sales," *Journal of Marketing Research*, 13 (November), 345-57.

Doyle, Peter and John Saunders (1990), "Multiproduct Advertising Budgeting," *Marketing Science*, 9 (Spring), 97-113.

Eastlack, Joseph O. and Ambar G. Rao (1986), "Modeling Response to Advertising and Pricing Changes for "V-8" Cocktail Vegetable Juice," *Marketing Science*, 5 (Summer), 245-59.

Feinberg, Fred (1992), "Pulsing Policies for Aggregate Advertising Models," *Marketing Science*, 11(Summer), 221-34.

Feinberg, Fred (2001), “On Continuous-Time Optimal Advertising Under S-Shaped Response,” *Management Science*, 47 (11), 1476-87.

Franses, Philip Hans (2005), “On the Use of Models for Policy Simulation in Marketing,” *Journal of Marketing Research*, 42, February, p. 4.

Gelfand, A. E., and A. F. M. Smith (1990), “Sampling-based Approaches to Calculating Marginal Densities,” *Journal of the American Statistical Association*, 85, 972-85.

Gelman, Andrew, John B. Carlin, Hal S. Stern and Donald B. Rubin (2004), *Bayesian Data Analysis*, Text in Statistical Science, Chapman & Hall/CRC, Boca Raton, Florida.

Gönül, Füsün F., Byung-Do Kim, and Mengze Shi (2000), “Mailing Smarter to Catalog Customers,” *Journal of Interactive Marketing*, 14 (2), 2-16.

------------------- and Mengze Shi (1998), “Optimal Mailing of Catalogs: A New Methodology Using Estimable Structural Dynamic Programming Models,” Management Science, 44 (9), 1249-62.

Green, Paul E. and Abba M. Krieger (1985), “Models and Heuristics for Product Line Selection,” *Marketing Science*, 4 (Winter), 1-19.

Greenberg, Allen and Charles Suttoni (1973), "Television Commercial Wearout," *Journal of Advertising Research*, 13 (5), 47-54.

Gupta, Sunil (1988), "Impact of Sales Promotions on When, What and How Much to Buy," *Journal of Marketing Research*, 25 (November), 342-55.

Gupta, Sunil (1991), "Stochastic Models of Interpurchase Time with Time-Dependent Covariates," *Journal of Marketing Research*, 28 (February), 1-15.

Hanssens, Dominique M. and Barton A. Weitz (1980), "The Effectiveness of Print Advertisements across Product Categories," *Journal of Marketing Research*, 17 (August), 294-306.

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

Keller, Kevin Lane (1991), "Memory and Evaluation Effects in Competitive Advertising Environments," *Journal of Consumer Research*, 17 (March), 463-76.

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), 317-32.

Leeflang, Peter S.H., Dick R. Wittink, Michel Wedel, and Philippe A. Naert (2000), *Building Models for Marketing Decisions*, International Series in Quantitative Marketing, Kluwar Academic Publishers, Dordrecht, The Netherlands.

Lenk, Peter J. and Ambar G. Rao (1995), “Transition Times: Distributions Arising from Time Heterogeneous Poisson Processes,” *Management Science*, 41 (7), 1117-29.

Little, John D. C. (1979), “Aggregate Advertising Model: The State of the Art,” *Operations Research*, 27 (4), 629-67.

Mahajan, Vijay, and Eitan Muller (1986), “Advertising Pulsing Policies for Generating Awareness of New Products,” *Marketing Science*, 5 (Spring), 86-106.

Mela, Carl F., Sunil Gupta, and Donald R. Lehmann (1997), “The Long Term Impact of Promotions and Advertising on Brand Choice,” *Journal of Marketing Research*, 34 (May), 248-61.

Mesak, H. I. and Zhang H. (2001), “Optimal Advertising Pulsation Policies: A Dynamic Programming Approach,” *Journal of the Operational Research Society*, 52, 1244-55

Neslin, Scott (2002) *Sales Promotion*, Marketing Science Institute, Cambridge, MA

Naik, Prasad A., Murali K. Mantrala, and Alan G. Sawyer (1998), “Planning Media Schedules in the Presence of Dynamic Advertising Quality,” *Marketing Science*, 17 (3), 214-35.

------------------- and Kalyan Raman (2003), "Understanding the Impact of Synergy in Multimedia Communication," *Journal of Marketing Research*, 40 (November), 375-88.

-- and Russell S. Winer (2005), "Planning Marketing-Mix Strategies in the Presence of Interaction Effects," *Marketing Science* 24(1), Winter, 25-34.

Park, Sehoon and Minhi Hahn (1991), "Pulsing in a Discrete Model of Advertising Competition," *Journal of Marketing Research*, 28 (November), 397-405.

Pieters, Rick G. M. and Tammo H. A. Bijmolt (1997), "Consumer Memory for Television Adverting: A Field Study for Duration, Serial Position, and Competition Effects," *Journal of Consumer Research*, 23 (March), 362-72.

Rossi, Peter E., Greg M. Allenby (2003), "Bayesian Statistics and Marketing," *Marketing Science***.** 22(3), 304-328

Russell, Gary J. (1988), "Recovering Measures of Advertising Carryover from Aggregate Data: The Role of Firm's Decision Behavior," *Marketing Science*, 7 (Summer), 252-70.

Schmittlein, David C. and D. G. Morrison (1983), "Prediction of Future Random Events with the Condensed Negative Binomial Distribution," *Journal of American Statistical Association*, 78, 448-56.

Sasieni, Maurice W. (1971), "Optimal Advertising Expenditures," *Management Science*, 18 (December), 64-72.

------------------------ (1989), "Optimal Advertising Strategies," *Marketing Science*, 8 (Fall), 358-70.

Simon, Hermann (1982), "ADPULS: An Advertising Model with Wearout and Pulsation," *Journal of Marketing Research*, 19 (August), 352-63.

Tellis, Gerald J., Rajesh K. Chandy, and Pattana Thaivanich (2000), "Which Ad Works, When, Where, and, How Often?: Modeling the Effects of Direct Television Advertising," *Journal of Marketing Research*, 37 (February), 32-46.

Thompson, Stephanie and Jack Neff (2002), "Kraft expands circ fro 'Food and Family'," *Advertising Age*, Vol. 73, Iss. 45, p. 41.

Villas-Boas J. Miguel (1993), "Predicting Advertising Policies in and Oligopoly: A Model and Empirical Test," *Marketing Science*, 12 (Winter), 88-102.

Wedel, Michel and Jie Zhang (2004), "Analyzing Brand Competition Across Subcategories," *Journal of Marketing Research*, 41 (November), 448-56.

Wentz Laurel (2006), "General Mills cooks up magazines for Hispanics," *Advertising Age*, Vol. 77, Iss 32, p. 25.

Weiss, Doyle L., Charles B. Weinberg, and Pierre M. Windal (1983), "The Effects of Serial Correlation and Data Aggregation on Advertising Measurement," *Journal of Marketing Research*, 20 (August), 268-79.

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

Zellner, Arnold (1987) *An Introduction to Bayesian Inference in Econometrics*, Robert E. Krieger Publishing Company, Inc., Malabar, Florida.

Zhang, Jie and Lakshman Krishnamurthi (2004) "Customizing Promotions in Online Stores," *Marketing Science*, 23 (Fall), 561-78.