MARKETING SCIENCE
Vol. 26, No. 2, March–April 2007, pp. 179–195
ISSN 0732-2399 | EISSN 1526-548X | 07 | 2602 | 0179
DOI 10.1287/mksc.1060.0208

Wearout Effects of Different Advertising Themes: A Dynamic Bayesian Model of the Advertising-Sales Relationship

Frank M. Bass,* Norris Bruce, Sumit Majumdar, B. P. S. Murthi
School of Management, The University of Texas at Dallas, Richardson, Texas 75083-0688
{mzjb@utdallas.edu, norris.bruce@utdallas.edu, majumdar@utdallas.edu, murthi@utdallas.edu}

Models of advertising response implicitly assume that the entire advertising budget is spent on disseminating one message. In practice, managers use different themes of advertising (for example, price advertisements versus product advertisements) and within each theme they employ different versions of an advertisement. In this study, we evaluate the dynamic effects of different themes of advertising that have been employed in a campaign. We develop a model that jointly considers the effects of wearout as well as that of forgetting in the context of an advertising campaign that employs five different advertising themes. We quantify the differential wearout effects across the different themes of advertising and examine the interaction effects between the different themes using a Bayesian dynamic linear model (DLM). Such a response model can help managers decide on the optimal allocation of resources across the portfolio of ads as well as better manage their scheduling. We develop a model to show how our response model parameters can be used to improve the effectiveness of advertising budget allocation across different themes. We find that a reallocation of resources across different themes according to our model results in a significant improvement in demand.

Key words: Bayesian dynamic linear models; Gibbs sampling aggregate advertising models; wearout effects; forgetting effects; copy effects; scheduling of ad copy
History: This paper was received June 30, 2005, and was with the authors 4 months for 1 revision; processed by Greg Allenby.

1. Introduction

Managers, policy makers, and researchers are interested in understanding the effect of advertising on demand. A large number of response models have been proposed in the literature linking advertising expenditures to sales or market shares. These studies have focused on the shape of the response function (linear or *S*-shaped), the dynamic effects of advertising (carry-over effects or wearout), and interaction effects with other marketing mix variables, consistent with the desirable properties of advertising response models in Little (1979).

In these response models, advertising expenditures are aggregated and there is an implicit assumption that advertising expenditure is spent on propagating one message or theme. In practice, firms concurrently run several themes in their advertising campaigns (for example, price and product advertisements). Each theme may have multiple versions, or executions, which are rotated over time. In effect, there is a portfolio of ads that is run in each week. Managers need to understand the individual effects as well as the interaction effects among the different versions of advertising on overall demand. Such an understanding then would allow managers to allocate their budgets over multiple themes more effectively. Most of the published research on the impact of advertising themes is categorized under copy research and these studies have employed experiments. They have focused on the effects of different copy on individual consumer attributes such as brand awareness and attitudes toward the brand. The literature that links the copy effects to sales or market share is limited to experimental studies and does not consider wearout effects (Aaker and Carmen 1982, Eastlack and Rao 1989).

The substantial research issues that we address in this study focus on understanding the effects of different themes of advertising on demand and on how to allocate a firm's resources across different themes to improve sales performance. The specific research questions pertinent to the study of the effectiveness of multiple themes in advertising are: How different are the wearout effects for the different themes of advertising? How can researchers accurately assess the magnitude of such wearout effects? What is the nature of interaction between the different themes of ads? For a given advertising budget, what is a more effective way to allocate resources across different advertising themes?

*Deceased.

Our response model extends the model by Naik et al. (1998) to multiple themes of advertising. They estimate a dynamic model of the effect of advertising on consumer awareness, rather than on demand, using Kalman filtering methods. Their model captured decay effects in the presence and absence of advertising and found that a gap in advertising has the effect of restoring ad quality. They also separate out the effects of two different types of wearout: copy wearout and repetition wearout. However, in their paper, they study the effectiveness of a single ad campaign with a single message and do not consider multiple themes of advertising which may be scheduled by the manager for the same product category. Our study generalizes their advertising model to account for multiple thematic executions of advertising. In addition, there are a few other notable differences between the two models. Naik et al. (1998) use brand awareness as the dependent variable while we use a measure of demand which is more managerially useful. We employ a different methodology, Bayesian estimation of a dynamic linear model (DLM) as in West and Harrison (1997).

Our paper makes three contributions. First, we extend earlier research on advertising by developing a generalized model of multiple themes of advertising which are aired concurrently in a given week. Such a model can allow managers to assess the wearout effects of different themes of advertising as well as consider the interaction effects between these different themes. Second, we capture the dynamic effects of different advertising themes as a function of wearout and forgetting. In other words, the parameters of the model are time varying and we are able to capture the dynamic path of these parameters. To do this, we employ Gibbs sampling in estimating the state space model of West and Harrison (1997). Two recent papers in marketing have also employed Gibbs sampling in a DLM framework (Neelamegham and Chintagunta 2004, Van Heerde et al. 2004). Leichty et al. (2005) employed Gibbs sampling to study the dynamic development of consumer preferences in a conjoint application. Finally, we develop a model to help managers conduct a what-if scenario analysis so that they can allocate their resources over different advertising themes more effectively.

Our data are obtained from a major telecom company that is a monopolist in its category. We have data on demand in terms of calling time, prices per minute of calling, and advertising spending for each of the themes that are defined by the company. A unique aspect of our data is that advertising is measured in gross rating points (GRPs) and not in dollars. One GRP represents advertising exposure to one percent of the population that owns television sets as defined by the ACNielsen TV ratings. There are two advantages of using GRPs instead of advertising dollars. First, GRPs provide a more accurate picture of advertising input than advertising expenditures since it is not clear how much advertising exposure can be purchased for a given budget. Second, most media buying is done in terms of GRPs and managers evaluate the effectiveness of their campaigns in terms of demand generated per GRP. The company classifies its advertising into five advertising themes: call stimulation ads, product offer ads, price offer ads, reconnect ads, and reassurance ads. We restrict our attention to these five themes.

The remainder of the paper is organized as follows. In §2, we provide a brief review of the relevant literature on estimation of advertising response with particular emphasis on wearout and dynamic effects. In §3, we present the details of our econometric model. In §4, we present the data and discuss the results of our estimation. Finally, we conclude with an overview of findings, the managerial implications, and the limitations of the study.

2. Literature Review

We present a brief overview of response models and review the literature on wearout effects and dynamic effects in greater detail. We also present arguments for studying the differences in wearout effects across multiple themes.

2.1. Response Models

In a seminal paper, Little (1979) stated that aggregate advertising response models should have the following desirable characteristics: (a) the effect of advertising should be nonlinear, (b) the models should capture the dynamic effects of wearout and forgetting, (c) models should consider the effect of competitive advertising, and (d) the ad effects could change over time due to changes in media and copy. Most of the models in the literature have been developed consistent with some, if not all, of these principles.

Early aggregate advertising response models linked advertising expenditures to sales or market share (Bass and Clarke 1972, Blattberg and Jeuland 1981, Hanssens et al. 1990) and considered the carry-over effects of advertising (Bass and Leone 1983, Broadbent 1984, Clarke 1976, Srinivasan and Weir 1988). These models used distributed lag models to capture the carry-over effect. Clarke (1976) showed that the magnitude of the effect of advertising and the duration of carry-over effects depended on the data interval used. The response models were then used to appropriate advertising dollars to maximize profits in both monopoly and oligopoly contexts (Simon 1965, Nerlove and Arrow 1962, Telser 1964, Palda 1964). These models used aggregate advertising expenditures and did not consider the effect of multiple themes. A good recent review of the advertising literature is in Vakratsas and Ambler (1999).

Another stream of research discussed the shape of the advertising response function: whether it is concave or *S*-shaped. Wittink (1977), Rao and Miller (1975), and recently Vakratsas et al. (2005) found evidence of an *S*-shaped function while Simon (1969) found no evidence for the *S*-shape. This was an important question because the theoretical models showed that the phenomenon of pulsing in advertising was related to the *S*-shape of the response function (Simon 1982, Mahajan and Muller 1986, Feinberg 1992). In a recent paper, Naik et al. (1998) show that pulsing can occur due to ad copy wearout, while Bronnenberg (1998) shows that pulsing can also occur in the context of a monopolist facing a Markovian sales response function.

2.2. Wearin and Wearout

Wearin refers to the positive effect on consumers who are exposed to an ad (Pechmann and Stewart 1990). The term wearout refers to the decay in advertising quality of an ad over time (Grass and Wallace 1969, Strong 1972, Calder and Sternthal 1980, Simon 1982). An ad is worn out if it either does not have any significant effect on consumers or has a negative effect. Both wearin and wearout effects depend on factors such as whether the ad was based on an emotional appeal or a rational appeal, whether the persuasion in the message was strong or weak, whether consumers were motivated or not to process the ad, and whether the level of competitive ads was high or low (Pechmann and Stewart 1990). The wearout effects may also depend on the change in ad copy. This is based on research that suggests that variations in copy improve the effectiveness or, specifically, recall of ads (Grass and Wallace 1969).

In a series of experiments to study of the effects of repetition of ads, Ray and Sawyer (1971a, b) found that the response functions for repetition varied across different measures (e.g., recall or intention), segments, brands, and type of advertising. They also studied the effect of different messages on repetition functions. Relevant to our study, they found that "grabber" ads were less effective over repetitions (i.e., had higher wearout) than "nongrabber" ads. MacInnis et al. (2002) find evidence of a significant positive relationship between ad repetition and sales when emotional ads are employed, but not for rational ads. They argue that one possible explanation of the above finding is that emotional ads have less rapid wearout.

Naik et al. (1998) model two sources of wearout—*repetition* wearout and *copy* wearout. When a customer is exposed to ads repeatedly, she can become bored, irritated, or simply lose interest as the benefits of processing the ad are perceived to be worthless (Berlyne 1970, Greyser 1973, Weilbacher 1970). This leads to repetition wearout, which depends on the amount of advertising that is done. Copy wearout, on the other hand, is the decay in advertising effectiveness due to the passage of time, which is independent of the amount of advertising. Such a decay may be the result of a change in consumers' conditions such as increased knowledge about product attributes over time (Calantone and Sawyer 1978). Other reasons for copy wearout include the imitation of an ad strategy by competing firms or by firms in other product categories (Axelrod 1980) and an increase in ad clutter (Corkindale and Newall 1978). A good discussion of these two effects is in Naik et al. (1998).

2.3. Forgetting

Another factor that affects quality dynamics is the effect of forgetting. Consumers tend to forget an ad when it is not being aired in a given period. Forgetting has a negative effect by reducing aggregate brand awareness (Mahajan et al. 1984). On the other hand, the literature also suggests that there is a rejuvenating effect of advertising when an ad is taken off the media (Grass and Wallace 1969, Greenberg and Suttoni 1973, Corkindale and Newall 1978, Naik et al. 1998). Grass and Wallace (1969) conclude that a period of no advertising enhances consumers' attention to ads. Similarly, Calder and Sternthal (1980) have found that the amount of cognitive responses increases when there is a break in advertising. The argument for improvement in ad quality, when it is not aired for a period of time, is that consumers may forget the particulars of a given ad and may feel that the ad is "*fresh*" when it is reintroduced. This suggests that as the period that an ad is pulled off the media is increased, there is a corresponding increase in forgetting and a consequent increase in the restoration of ad quality (Corkindale and Newall 1978).

2.4. Differential Effects of Wearout Across Themes

As stated earlier, wearout is affected by several factors such as the type of appeal, level of competitive advertising, and strength of persuasion. Experimental evidence shows that emotional ads wearout more slowly than ads based on nonemotional or rational appeals (Hitchon et al. 1988). In this laboratory study, the dependent variables were attitude toward the ad, brand attitudes, and purchase intentions. The authors found that unemotional ads exhibited faster wearout in all of the three dependent variables when compared to emotional ads. This could be due to the fact that ads with emotional images elicit imagery processing while verbal arguments elicit cognitive processing (MacInnis and Price 1987). Silk and Vavra (1974) and Ray and Sawyer (1971b) also suggest that "*soft sell*" ads, which use emotional images, wear out slower than "*hard sell*" ads, which are based on verbal arguments. These arguments highlight the need for studying the differential wearout effects of different themes of advertising and their effect on demand.

2.5. Need for Time-Varying Coefficients

Advertising effects accumulate over time, hence it is reasonable to expect that the response coefficients will also vary over time. Researchers have modeled the phenomenon of time-varying parameters using data from different time periods and estimating a separate coefficient for each time period (Winer 1979, Bronnenberg et al. 2000, Mela et al. 1997). Since only a part of the data is used in each estimation, these estimates are likely to be inefficient.

Other researchers have employed a random coefficients approach (Jedidi et al. 1999) where the parameter is assumed to be distributed according to a probability distribution over time. In these models, the variance of the distribution is estimated; thus researchers are able to control for the time-varying nature of the response parameters but are unable to recover the parameter paths over time. The model used in our research, the dynamic linear model (DLM), is able to estimate the dynamic path of the response coefficient over time and the estimates are efficient.

In a recent paper, Dubé and Manchanda (2005) study the differences in dynamics of marketing mix (price and advertising) across several markets. They conclude that a firm's current and past advertising has a larger effect on its own demand, especially in larger markets. They also conclude that competitors' advertising has a much smaller effect on a firm's demand than the firm's own advertising. This work suggests that ignoring competitive advertising while modeling a firm's advertising-to-sales relationship is unlikely to cause serious error.

2.6. Need for Interactions

The interaction effects between advertising and other marketing mix variables, especially price, have been well-documented (Eskin and Baron 1977, Wittink 1977, Krishnamurthi and Raj 1985, Winer and Moore 1989). Therefore, it is necessary to model the interaction between the different themes of advertising as well. A recent paper by Naik et al. (2005) makes a strong case for modeling interactions between advertising and promotions in developing demand models in a competitive environment. In sum, our paper extends the literature on advertising response models by considering wearout effects, interaction effects, and dynamic effects of different themes of advertising while being consistent with prior research guidelines.

Montgomery et al. (2005) use a survey to provide evidence that managers, when making decisions, place considerably less weight on competitive reactions and strategic competitive reasoning.

3. Model

We develop a model of advertising that links demand to the advertising expenditure that is incurred in airing multiple themes of advertising in a given week. Since advertising effects last longer than a week, we can assume that multiple themes are being aired concurrently. The model should consider the dynamic effects of wearout and forgetting as described earlier.

We begin by generalizing the model proposed by Naik et al. (1998), who estimate a model that links awareness to the advertising expenditure data. Though they estimated a number of advertising models that have been proposed in the marketing literature such as Nerlove and Arrow (1962), Vidale and Wolfe (1957), Little (1975), Blattberg and Golanty (1978), and Blackburn and Clancy (1982), they find that the Nerlove-Arrow model provides them the best fit. Given this information, we focus our attention in this paper on the Nerlove-Arrow model.

In the Nerlove-Arrow model, the rate of change in goodwill $G(t)$ is a function of the advertising expenditure per week $A(t)$ that can be measured either in terms of gross rating points (GRP) or in dollars. Specifically,

$$\frac{dG(t)}{dt} = qA(t) - \delta G(t), \tag{1}$$

where

$q =$ effectiveness of ad spending (assumed constant, set equal to one in Nerlove-Arrow).
$\delta =$ rate of decay of goodwill due to forgetting.
$A(t) =$ advertising spending in each time period t.
$G(t) =$ goodwill in each time period.

More generally, one can use a function of advertising expenditure $f(A(t))$, which need not be linear as in the Nerlove-Arrow model.

Naik et al. (1998) extend the above model by making the advertising effectiveness parameter q to be time dependent. In other words, q is not assumed to be constant over time, but is a function of time as well as advertising expenditure. The rationale for such an approach is that it allows the capture of different types of wearout effects—copy wearout and repetition wearout. Copy wearout refers to the decay in advertising effectiveness over time due to the message becoming less effective, and this effect is assumed to be independent of the amount of advertising. Naik et al. (1998) point out that copy wearout can occur due to several reasons: (1) the consumer becomes more knowledgeable about the product attributes, (2) the amount of competitive advertising in response to this firm's advertising can reduce attention, or (3) other firms imitate the advertising style. Repetition wearout, on the other hand, depends on the amount of advertising expenditure and occurs because consumers get bored of seeing the same ad or perceive the value of processing the ad a second time as being small. The greater the amount of advertising, the greater is the repetition wearout effect.

To describe the evolution of the effectiveness parameter q over time, Naik et al. (1998) propose a differential equation as follows:

$$\frac{dq}{dt} = -a(A)q + (1 - I(A))\delta(1 - q) \tag{2}$$

where

$$a(A) = c + wA(t), \tag{3}$$

c is the copy wearout parameter and w is the repetition wearout parameter.

In the above equations, there are two types of effects. When the advertising is "*on*" in a given time period, the indicator function $I(A)$ takes the value $= 1$ and the second term on the RHS of Equation (2) becomes zero. The effect of advertising is captured by copy wearout parameter c and repetition wearout parameter w in Equation (3). Note that repetition wearout depends on advertising level $A(t)$. When advertising is "*off*," there is a rejuvenating effect of advertising that is captured by the second term on the RHS of Equation (2).

We now generalize the Nerlove-Arrow and Naik models to accommodate multiple ad themes. A different combination of ad themes is aired in each week. We model the ad spending rate in Equation (1) as an additive function of ad spending on individual themes. The modified equation of the rate of change in goodwill is given by a generalization of the Nerlove-Arrow model for advertising themes:

$$\frac{dG}{dt} = \sum_{i=1}^{m} q_i \left(g(A_i) + \lambda_i \sum_{\substack{j=1 \\ j \neq i}}^{m} h(A_i, A_j) \right) - \delta G, \tag{4}$$

where q_i is the effectiveness of ad theme i, $g(A_i)$ is a function of the advertising expenditure for theme i, m is the number of ad themes, G is the goodwill, and A_i is the advertising expenditure for each ad theme. Note that both goodwill and advertising vary over time. Because of availability of data on several ad themes, we are able to understand the interaction effects among the different themes. The term

$$\lambda_i \sum_{\substack{j=1 \\ j \neq i}}^{m} h(A_i, A_j)$$

is an interaction effect for the ith theme. We estimate a separate interaction effects parameter (λ_i) for each theme, thus allowing for possible asymmetric effects. It is an overall measure of how the ith theme interacts with all other concurrent advertising themes at time t. We suppress the notation for time for ease of exposition. One can potentially estimate interaction parameters for all pairwise combinations of ad themes, but we estimate only one coefficient for each theme for the sake of parsimony. Furthermore, we have no theory to explain the nature of pairwise interactions, which could depend on the nature of ad themes and their execution.

In the estimation, we use $g(A_i) = \ln(1 + A_i)$ and $h(A_i, A_j) = \ln(1 + A_i)\ln(1 + A_j)$. The justification for use of the semilog specification has been presented in Doyle and Saunders (1990). They rule out linear and exponential functional forms as they do not exhibit diminishing returns, which is a characteristic of advertising response models (Simon 1970). Polynomial functions exhibit supersaturation and do not fare well in optimization (Doyle and Saunders 1990). The semilog specification has also been favored by Lambin (1969) and Carroll et al. (1979). Jagpal et al. (1979) also recommend using a log-log specification and consider the interaction term as the product of two $\ln(A_i)$ terms as we have done.

The change in ad effectiveness q_i is given by the following equation:

$$\frac{dq_i}{dt} = -a(A_i)q_i + \delta\{(1 - I(A_i))\}(1 - q_i), \quad i = 1, 2, \ldots, m \tag{5}$$

where

$$a(A_i) = c_i + w_i A_{it}. \tag{6}$$

Note that there are five different equations for Equation (5), one for each advertising theme. Further, for tractability we assume that the rate of change in ad quality for the ad themes (dq_i/dt) are independent of each other. There is some support for the assumption of independence in the study by Blair and Rabuck (1998). Based on an analysis of over 500 case studies they conclude that commercials within a campaign wear out independently of their pool partners. Even though there may be similarities between some ad themes which could be modeled, such a dependency would complicate the estimation of the dynamic parameters in our model. We therefore leave the modeling of such dependencies between ad themes for future research.

Thus, our model allows for the estimation of dynamic effects of advertising themes in the presence of wearout effects. We control for cross-sectional heterogeneity between the ad themes by estimating separate dynamic parameters for each ad theme. We employ the Gibbs sampler for estimation of our parameters, as in West and Harrison (1997). There are two advantages of using a DLM model. First, we are able to provide an understanding of how the response parameters themselves change over time. To understand the differences in advertising effectiveness of two different ad themes, it is not enough to observe differences in mean parameter values because even if the average effect of a particular response coefficient may be the same across two advertising themes, their parameter paths may be substantially different over time. Other researchers have employed a random coefficients specification to model time-varying parameters (Jedidi et al. 1999). Such models can control for the time-varying nature of parameters but do not provide an estimate of the parameter at a given point in time. DLM also handles missing observations in a trivial manner as no updating takes place.

Second, advertising response may be nonstationary. In time-series analysis, researchers filter the data (say, by taking first- or higher-order differences) in order to make the series stationary. West and Harrison (1997) point out that these filtering methods affect the interpretation of the model by confounding different model components. Further sources of nonstationarity that deviate from the process implied by the filter are not captured. They suggest that DLM presents a better method for handling nonstationarity.

4. Data and Estimation

4.1. Data

We obtained data for a major telecommunications services company in which the demand for residential telephone services has been measured in two ways: minutes of call time and number of calls. We have used the total call time in millions of hours as our dependent variable. We aggregated demand for three call types—local, regional, and national—each of which are classified as weekday and weekend calls. We have not considered international calls. The covariates that are available to model demand are the average price per minute of a call, the number of land lines available, and competitive advertising. We compute the weighted average price per minute over the different categories that we considered since the rates are different for these categories. In addition, we have data on the number of lines available to account for the increase in capacity over the 114 weeks under study. Note that the telecommunications company is a monopoly in the land line telecommunications business and competition exists only in the wireless markets. We group all competitive advertising expenditures into one category.

The company classifies its advertising into five themes—call stimulation ads, product offer ads, price offer ads, reconnect ads, and reassurance ads. Advertising expenditure for each theme of advertising is measured in terms of gross rating points (GRP). Advertising GRPs of different themes of advertising are available for a period of 114 weeks between 1995 and 1997. In Figure 1, we provide plots of GRPs for

Figure 1 Plot of GRPs for Different Themes Over Time

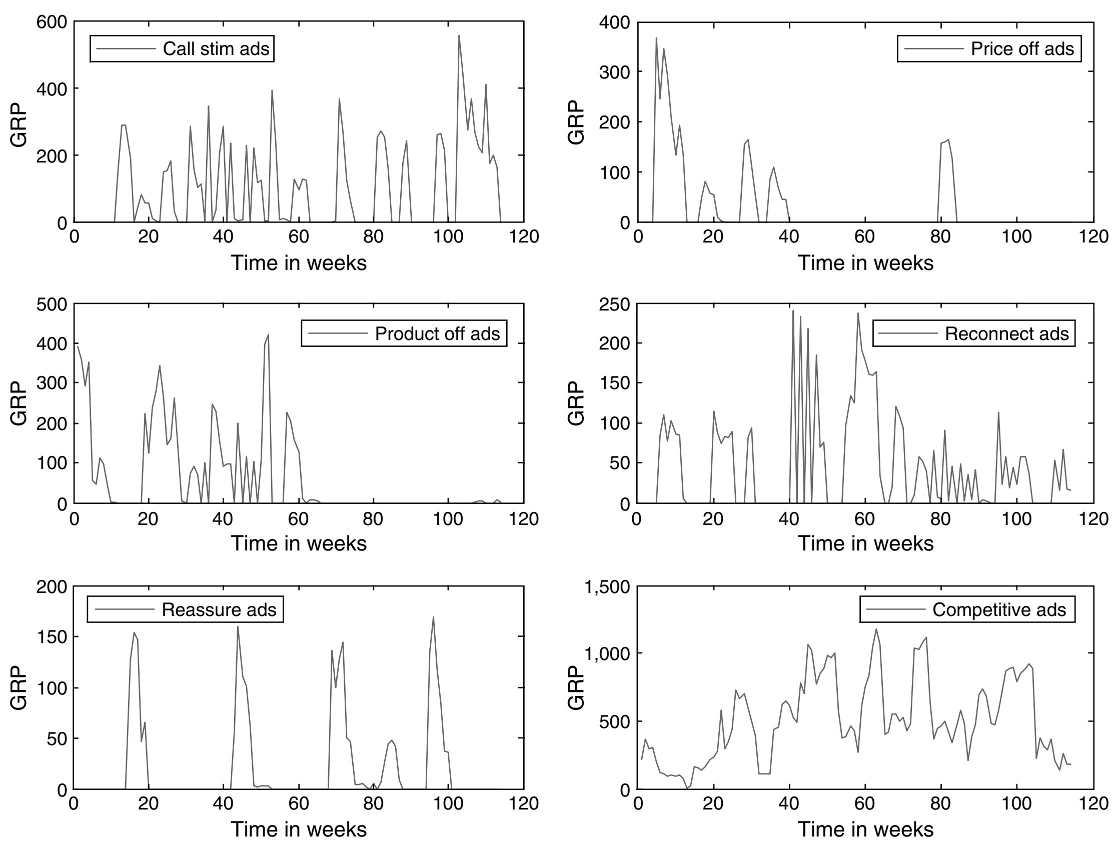
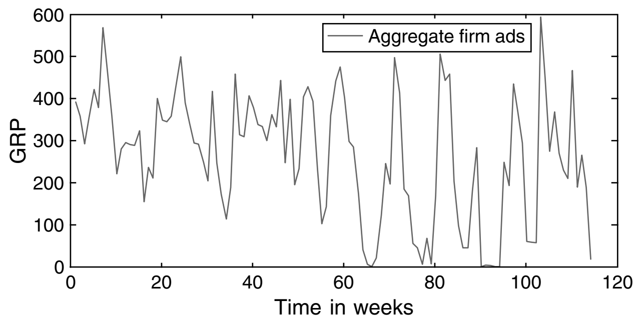


Figure 2 Plot of Total GRP Over Time



different advertising themes as well as for total GRP of competitive ads over the 114-week period. The total amount of GRP across all themes for the firm is given in Figure 2. Descriptive statistics are given in Tables 1 and 2.

4.2. Model Estimation

We model the demand for residential telephone services (y_t) at time t as a function of goodwill (G_t), a vector $\mathbf{X}_t$ with variables price, number of lines, and competitive advertising, and mean-zero normally distributed error ε_t:

$$y_t = G_t + \boldsymbol{\beta}\mathbf{X}_t + \varepsilon_t \quad \text{where } \varepsilon_t \sim N(0, \sigma_\varepsilon^2). \tag{7}$$

This unobservable error (ε_t) could be related to demand factors such as the growth in mobile telephone service or the availability of the Internet. We assume these omitted demand factors are perceived by the firm and can influence price, suggesting a potential endogeneity problem (Villas-Boas and Winer 1999). Therefore, we use instruments (W) to address this problem. We used retail price index, number of households, consumer sentiment, and household spending as instruments in addition to competitive advertising and number of lines:

$$p_t = p_t(W; \alpha) + \eta_t. \tag{8}$$

Another source of endogeneity bias could be due to the assumption of advertising GRP as an exogenous variable. If firms choose the level of advertising strategically (say, in response to competitors' advertising), then these variables are endogenous. This potential endogeneity bias could be mitigated if we had good instruments in our data. We leave the study of endoeneity bias as an issue for future research. We expect that the potential bias in estimates could affect the solution to our allocation model.

To estimate the parameters of our model, we convert the above system of Equations (4)–(6) to discrete

We are grateful to an anonymous reviewer for bringing this issue to our attention.

Table 1 Descriptive Statistics of Main Variables

	Mean gross rating points	Standard deviation
Call stimulation ads	100.96	128.59
Price ads	31.68	73.21
Product ads	63.87	107.77
Reconnect ads	45.62	61.33
Reassurance ads	21.25	44.01
Competition ads	505.82	
Mean call volume (hours)	16.42 million	
Mean weighted price per minute	2.46	
Mean line capacity (number of lines)	22.63 million	

time and rewrite the generalized model in more formal, state space notation:

$$y_t = \underbrace{[1 \quad 0 \quad \cdots \quad 0]}_{\mathbf{F}_t} \underbrace{\begin{bmatrix} G_t \\ q_{1t} \\ \cdots \\ q_{mt} \end{bmatrix}}_{\boldsymbol{\Phi}_t} + \boldsymbol{\beta}'\mathbf{X}_t + \varepsilon_t \tag{9}$$

$$\underbrace{\begin{bmatrix} G_t \\ q_{1t} \\ \cdots \\ q_{mt} \end{bmatrix}}_{\boldsymbol{\Phi}_t} = \underbrace{\begin{bmatrix} (1-\delta) & \bar{g}(A_{1t}) & \cdots & \bar{g}(A_{mt}) \\ 0 & \begin{matrix}(1-a(A_{1t})) \\ -\delta(1-I(A_{1t}))\end{matrix} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \begin{matrix}(1-a(A_{mt})) \\ -\delta(1-I(A_{mt}))\end{matrix} \end{bmatrix}}_{\mathbf{H}_t} \underbrace{\begin{bmatrix} G_{t-1} \\ q_{1t-1} \\ \cdots \\ q_{mt-1} \end{bmatrix}}_{\boldsymbol{\Phi}_{t-1}}$$

$$+ \underbrace{\begin{bmatrix} 0 \\ \delta(1-I(A_{it})) \\ \cdots \\ \delta(1-I(A_{mt})) \end{bmatrix}}_{\mathbf{u}_t} + \underbrace{\begin{bmatrix} w_{0t} \\ w_{1t} \\ \cdots \\ w_{mt} \end{bmatrix}}_{\mathbf{w}_t}, \tag{10}$$

where

$$\bar{g}(A_{it}) = g(A_{it}) + \lambda_i \sum_{\substack{j=1 \\ j\neq i}}^{m} h(A_{it}, A_{jt}), \quad i = 1, \ldots, m.$$

Note that, as stated earlier, in the estimation we use $g(A_{it}) = \ln(1 + A_{it})$ and $h(A_{it}, A_{jt}) = \ln(1 + A_{it}) \cdot \ln(1 + A_{jt})$.

As shown above, $\boldsymbol{\Phi}_t$ is the state vector whose first element is goodwill and remaining elements are ($m = 5$) advertising qualities, one for each theme during time t. The $(m+1) \times (m+1)$ transition matrix $\mathbf{H}_t$ captures the time-varying effects of ad spending, wearout, and forgetting on the portfolio of advertising themes and goodwill across time. The constant vector $\mathbf{F}_t$ reflects the impact of goodwill on telephone usage. Error terms $\mathbf{w}_t$ and ε_t are assumed to be mean-zero independent normals. Thus, we can rewrite Equations (9) and (10) in a more compact notation to obtain the

Table 2 Correlation Between GRPs of Different Themes

	Price ads	Product ads	Reconnect ads	Reassurance ads	Competitive ads
Call stimulation ads	−0.137 (ns)	−0.242**	−0.269**	−0.014 (ns)	−0.02 (ns)
Price ads		−0.038 (ns)	0.037 (ns)	−0.157*	−0.35*
Product ads			−0.029 (ns)	−0.175*	−0.024 (ns)
Reconnect ads				0.0001 (ns)	0.186*
Reassurance ads					0.119 (ns)

*Significant at $p < 0.10$.
**Significant at $p < 0.01$; ns indicates nonsignificant.

standard Bayesian dynamic linear model (DLM) of West and Harrison (1997):

$$y_t = \mathbf{F}_t\boldsymbol{\Phi}_t + \boldsymbol{\beta}\mathbf{X}_t + \varepsilon_t, \tag{11}$$

$$\boldsymbol{\Phi}_t = \mathbf{H}_t\boldsymbol{\Phi}_{t-1} + \mathbf{u}_t + \mathbf{w}_t,$$

where

$$\varepsilon_t \sim N[0, \sigma_\varepsilon^2],\ \mathbf{w}_t \sim N[0, \mathbf{W}]. \tag{12}$$

Gathering the parameters of the model together in a tuple constitutes the DLM specification $\{\mathbf{F}_t, \mathbf{H}_t, \sigma_\varepsilon^2, \mathbf{W}\}$. The model is then completed with specifications of: (i) prior distributions for the unobserved elements of the transition matrix $\mathbf{H}_t$, δ, $\{c_i, w_i, \lambda_i\}_{i=1}^m$, (ii) prior distributions for $\boldsymbol{\beta}$, the coefficient of other independent measures that affect the demand for telephone services, (iii) prior distributions for unobserved observation and system variances σ_ε^2 and $\mathbf{W}$, and (iv) initial guess for the state vector of containing goodwill and advertising quality $\boldsymbol{\Phi}_0$ at time $t = 0$ based on initial information denoted by $\mathbf{D}_0$. We will assume standard conjugate forms for all priors in our analysis. Prior specifications are chosen to allow the data to dominate the results.

While many state space methods (Harvey 1994) rely on maximum likelihood estimation (MLE) for estimation, the DLM is based on Bayesian estimation. West and Harrison (1997) suggest the use of traditional Monte Carlo Markov chain (MCMC) techniques such as Gibbs sampling. There are several reasons why the DLM is particularly suited to studying time-varying advertising effects. Relative to the static Bayesian models seen in the marketing literature (Allenby and Rossi 2003), the DLM offers improved estimation via adaptation and Bayesian learning. Moreover, with suitable informative priors, relatively accurate forecasts can be produced from series too short for purely data driven (frequentist) analyses.

We estimate the above system of discrete Equations (9) and (10) by specifying priors for the model parameters and using MCMC simulation of the full posterior based on the entire data series $t = 1, 2, \ldots, T$. Let $\boldsymbol{\Phi}^T = \{\boldsymbol{\Phi}_0, \boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \ldots, \boldsymbol{\Phi}_T\}$ and $\mathbf{y}^T = \{y_1, y_2, \ldots, y_T\}$ be the state parameters and telephone usage over the entire data set. Let $\boldsymbol{\Lambda} = \{\delta, \{c_i, w_i, \lambda_i\}_{i=1}^m\}$ be a vector of forgetting, wearout, and interaction parameters defined in $\mathbf{H}_t$, and $\boldsymbol{\beta}$ the coefficient for other explanatory variables that influence telephone usage. Now assume that the prior on σ_ε^2 is inverse gamma and the prior on $\mathbf{W}$ is independent inverse Wishart. Then, by using a direct Gibbs sampling approach (Gelfand and Smith 1990), we can compute the complete joint posterior $p(\boldsymbol{\Phi}^T, \mathbf{W}, \sigma_\varepsilon^2\boldsymbol{\Lambda}, \boldsymbol{\beta} \mid \mathbf{y}^T)$, iteratively resampling conditional posteriors $p(\boldsymbol{\Phi}^T \mid \mathbf{y}^T, \sigma_\varepsilon^2, \mathbf{W}, \boldsymbol{\Lambda}, \boldsymbol{\beta})$ and $p(\mathbf{W}, \sigma_\varepsilon^2, \boldsymbol{\Lambda}, \boldsymbol{\beta} \mid \boldsymbol{\Phi}_t, \mathbf{y}^T)$ (see Appendix A for details).

4.3. Model Identification

To show identification in our DLM, we need to investigate whether enough *prior structure* exists in the system of Equations (11) and (12). First, we need to construct an observationally equivalent system $M_1 = \{\mathbf{F}_{1t}, \mathbf{H}_{1t}, \sigma_\varepsilon^2, \mathbf{W}_1\}$, where *observability* (Crassidis and Junkins 2004) relates to our ability to recover all state vectors $\boldsymbol{\Phi}^T = \{\boldsymbol{\Phi}_0, \boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \ldots, \boldsymbol{\Phi}_T\}$ given observations $\mathbf{y}^T = \{y_1, y_2, \ldots, y_T\}$. An *equivalent* system is achieved by substituting the linear transformation $\boldsymbol{\Phi}_{1t} = \mathbf{L}\boldsymbol{\Phi}_t$ into Equations (11) and (12), where $\mathbf{L}_t$ is a $(m+1) \times (m+1)$ nonsingular matrix. We obtain $\mathbf{F}_{1t} = \mathbf{F}_t\mathbf{L}^{-1}$, $\mathbf{H}_{1t} = \mathbf{L}\mathbf{H}_t\mathbf{L}^{-1}$, $\mathbf{u}_{1t} = \mathbf{L}\mathbf{u}_t$, and $\mathbf{W}_1 = \mathbf{L}\mathbf{W}\mathbf{L}'$. Under this transformation, the model M_1 is indistinguishable from our system described in Equations (11) and (12). Thus, for identification it is sufficient to show that enough *prior structure* exists in $\mathbf{F}_t$ and $\mathbf{H}_t$ such that all transformations are precluded except $\mathbf{L} = \mathbf{I}$ (the identity matrix). It turns out that $\mathbf{L} = \mathbf{I}$ for the Naik et al. (1998) and our model (see Appendix B for a proof).

Finally, for the incorporation of price endogeneity, we take a two-step approach. In the first step, we estimate a Bayesian regression model with all the exogenous variables and instruments as independent variables. In the second step, we replace prices in the

For comparison, we also implemented a fully Bayesian instrumental variable procedure given in Rossi et al. (2005). Their approach is different from that in Lancaster (2004). An important distinction between these two treatments is the specification of the prior on the covariance matrix of the reduced form errors. This matrix is a function of the coefficient of the endogenous variable. Lancaster (2004) assumes that the prior is independent of the coefficient of the endogenous variable while Rossi et al. (2005) do not, which we believe is more reasonable. We find that our substantive results were unchanged.

DLM observation equation with their respective predicted values obtained from the first stage estimation. These predicted values are obtained from the posterior predictive distribution, consistent with our Bayesian framework.

5. Results

We report the results of our estimation in Tables 3 to 9. In Table 7, we provide values of the log Bayes factors for comparing the model fits across alternate models. In Tables 4 and 5, we provide estimates of our complete model with interactions between different themes. We also estimate a model in which there are no interaction effects between the different ad themes (Table 7). To compare the results, we estimate another model in which we aggregate the advertising expenditures of different themes and treat the demand as a function of total advertising GRPs (Table 8).

To compare model fit across two models that differ in parameters, we employ Bayes factors or log Bayes factors (West and Harrison 1997, Congdon 2001). A log Bayes factor of one or greater indicates evidence in favor of our model. From Table 3(a), we see that a model with no advertising interactions is rejected in favor of a model with advertising interactions. Similarly, a model that considers aggregate advertising GRPs across the different themes as a dependent variable is also rejected in favor of our proposed model; this highlights the value of modeling different responses for different advertising themes. We also reject the linear and square root functional forms for $g(A_{it})$ in favor of the $\ln(A_{it})$ specification.

We also compare the predictive performance of alternate models and compute the mean absolute deviation (MAD) and mean square error (MSE) between predictions of demand from the model and the actual demand. In Tables 3(b) and 3(c), we compute MAD and MSE using data from all 114 weeks and use the parameters to predict demand for 1 to 114 weeks and 51 to 114 weeks. The MAD and MSE values in both sections are the lowest for the full model with interactions. We also provide out-of-sample forecast performance of alternate models in Table 3(d). We estimate Equations (9) and (10) using two sample timeframes ($t = 1$:60, and $t = 1$:100) and generate out-of-sample (step ahead) forecasts for 10 time periods $k = 1, 2, 3, \ldots, 10$. We report MAPE (mean absolute percentage error) across the $k = 10$ time periods for all model forecasts. The evidence clearly supports the predictive performance of our proposed model.

In Table 7, we present the estimates of our model and we see that the effect of price per minute on demand is negative. To obtain an understanding of the significance level, we report the 95% HPDI (highest probability density interval). We note that price per minute has a significant effect on demand at 95% confidence level. Further, capacity as measured by the number of lines also has a significant positive effect on demand, while competitive advertising has a negative but nonsignificant effect on demand. The lack of an effect due to competition may be due to the fact the firm is dominant in its category. These effects are consistent with our expectations.

The forgetting rate δ is 0.034, significant and consistent with values obtained in earlier studies (Naik et al. 1998). This means that when an ad is taken off consumers tend to forget the message and this could lead to an enhancement in quality of the ad when it is aired again after some time. The copy wearout effects c are all positive and significant. The coefficients are different for different themes and range from a low value of 0.1625 for reassurance ads to 0.5725 for price offer ads. This suggests that price offer ads have a higher

Table 3 Comparison of Current Specification with Alternate Specifications

(a) Log Bayes Factors

Specification	Description	Log Bayes factor
Model 2	No advertising interactions	98.767
Model 3	Aggregate GRPs	87.712
Model 4	Square root GRP	21.565
Model 5	Linear GRP	360.19

Note. $T = 1$: 116 weeks.

(b) Predictive Performance

Specification	Description	MAD	MSE
Model 1	Full model	0.4194	0.3499
Model 2	No advertising interactions	0.9837	1.4942
Model 3	Aggregate GRPs	0.9072	1.2871
Model 4	Square Root GRP	0.7538	0.9082
Model 5	Linear GRP	1.4564	3.0890

(c) Predictive Performance

Specification	Description	MAD	MSE
Model 1	Full model	0.4752	0.4727
Model 2	No advertising interactions	1.2390	2.2029
Model 3	Aggregate GRPs	1.0398	1.6736
Model 4	Square Root GRP	0.8975	1.2051
Model 5	Linear GRP	1.4017	3.0049

Note. $T = 50$: 116 weeks.

(d) Model Forecast Performance

Specification	Description	MAPE[1]	MAPE[2]
Model 1	Full model	6.31	5.92
Model 2	No advertising interactions	18.94	29.63
Model 3	Aggregate GRPs	51.92	13.08
Model 4	Square root GRP	11.86	27.53
Model 5	Linear GRP	253.96	238.87

[1]Sample $t = 1$:60.
[2]Sample $t = 1$:100.

Table 4 Estimates by Message Type (with Interactions)

Parameters	Estimate	Standard deviation	95% HPDI	
Price per minute	−11.9171	6.3939	−22.6110	−1.4258
Competitive advertising	−0.0467	0.1434	−0.2860	0.1844
Capacity (number of lines)	1.1192	0.5879	0.1347	2.0848
Forgetting rate δ	0.0344	0.0089	0.0236	0.0469
Initial goodwill G_0	14.326	0.9774	12.833	16.014
Observation variance σ_ε	0.1280	0.0287	0.0870	0.1796
System variance (goodwill) $\sigma_{\varepsilon 1}$	0.0796	0.0103	0.0642	0.0979
Call stimulation				
Copy wearout c	0.2494	0.0893	0.1421	0.3649
Repetition wearout w	−0.0329	0.0222	−0.0612	−0.0053
Initial quality q_{10}	0.1214	0.0627	0.0252	0.2316
System variance (ad quality) $\sigma_{\varepsilon 2}$	0.0020	2.5599e−4	0.0016	0.0024
Price offer				
Copy wearout c	0.5725	0.1947	0.3264	0.8363
Repetition wearout w	−0.1084	0.0393	−0.1619	−0.0597
Initial quality q_{20}	0.0604	0.0263	0.0220	0.1078
System variance (ad quality) $\sigma_{\varepsilon 3}$	1.984e−4	2.5863e−5	1.6008e−4	2.4405e−4
Product offer				
Copy wearout c	0.4041	0.1279	0.2420	0.5725
Repetition wearout w	−0.0666	0.0289	−0.1037	−0.0294
Initial quality q_{30}	0.0809	0.0316	0.0299	0.1344
System variance (ad quality) $\sigma_{\varepsilon 4}$	7.9286e−4	1.0349e−4	6.3891e−4	9.7657e−4
Reconnection				
Copy wearout c	0.2275	0.0644	0.1577	0.3087
Repetition wearout w	−0.0488	0.0143	−0.0670	−0.0330
Initial quality q_{40}	0.1370	0.0459	0.0678	0.2182
System variance (ad quality) $\sigma_{\varepsilon 5}$	7.9796e−5	1.0335e−5	6.4505e−5	9.7978e−5
Reassurance				
Copy wearout c	0.1625	0.0606	0.1064	0.2315
Repetition wearout w	−0.0432	0.0157	−0.0617	−0.0274
Initial quality q_{50}	0.1800	0.0595	0.0839	0.2816
System variance (ad quality) $\sigma_{\varepsilon 6}$	1.5915e−4	2.0762e−5	1.2836e−4	1.947e−4

rate of copy wearout than that of reassurance ads. We provide a discussion of these parameters later.

We find that the repetition wearout effects are negative and significantly different from zero. For instance, the repetition wearout effect for call stimulation ad is −0.0329. This suggests that there is no repetition wearout, but in fact there is wearin. It is important to note that Naik et al. (1998) obtain positive and significant effects for repetition wearout while we get significant negative coefficients. This may be explained by the fact that the firm employed different themes and multiple executions within each theme. This rotation of ads across different themes could have the effect of keeping the ads fresh and may explain the wearin effect of advertising. This is consistent with Lodish et al. (1995), in which they conducted a meta-analysis of a large number of split cable television experiments and concluded that an increase in sales due to an increase in media weight is more likely when copy strategy is changed. So while repeating the same ad might cause wearout, having different executions of the same message may contribute to a wearin. Research in laboratory experiments also indicates that repeated ads using multiple executions lead to better recall (Unnava and Burnkrant 1991). Our results on wearout differ from Naik et al. (1998) and demonstrate the importance of considering the effect of different themes in obtaining correct wearout effects.

Table 5 Analysis of Wearout Parameters

	No. of executions	Copy wearout parameter	Repetition wearout parameter
Rational appeal			
Price offer ads	60	0.5725	−0.1084
Product ads	22	0.4041	−0.0666
Average	41	0.4883	−0.0875
Emotional appeal			
Call stimulation ads	43	0.2494	−0.0329
Reassurance ads	17	0.1625	−0.0432
Reconnect ads	16	0.2275	−0.0488
Average	25.3	0.2131	−0.0416

To understand the reason for differences in copy effects between different themes, we draw on the findings in the review by Pechmann and Stewart (1990). They suggest that emotional ads and ads with greater emotional imagery wear out slower than ads without emotional persuasion. We conduct an informal analysis of this argument as we do not have information about the actual content. We speculate that price and product offer ads have a more rational appeal and we suspect that call stimulation, reassurance, and reconnect ads have a greater emotional appeal. Based on this classification, we see in Table 5 that rational ads have larger copy wearout parameters than emotional ads. These findings are consistent with the research by Hitchon et al. (1988) and the verbal arguments in Silk and Vavra (1974) and Ray and Sawyer (1971) which suggest that "*hard sell*" ads wear out faster than "*soft sell*" ads.

Further, the literature review posits that wearout can be mitigated by using a number of different copy executions. We show the number of executions for each theme and the repetition wearout parameters in Table 5. Within the category of rational ads, we see that as the number of executions increases, the repetition wearout decreases. We do not find the same

Table 6 Advertising Interactions

Copy interactions	Estimate	Standard deviation	95% HPDI	
Call stimulation	−0.0773	0.0268	−0.1194	−0.0314
Price offer	−0.0162	0.0548	−0.0957	0.0818
Product offer	−0.0384	0.0362	−0.0949	0.0228
Reconnection	−0.0846	0.0273	−0.1217	−0.0372
Reassurance	−0.0969	0.0228	−0.1277	−0.0595

pattern within the category of emotional appeal ads; however, if we take the average number of executions across rational and emotional ads, there is a link between the number of copy executions and wearin. As the number of copy executions increases, the wearin is higher (i.e., wearout is lower). While these conclusions are not definitive evidence, they are consistent with what researchers have observed in laboratory settings and provide validity to our model and estimation. Note that the above analysis is based on our ad hoc classification of ads as either rational or emotional.

In Table 6, we present the interaction effects of the various themes. We find that all the interaction coefficients are negative, and the interactions of call stimulation, reconnection, and reassurance ads are statistically significantly different from zero. Note that in our model we assume that the interaction of, say, a call stimulation ad with all other ads has the same coefficient. The negative parameters suggest that the interaction between the different themes mitigates the goodwill generated by the ad campaign. There is considerable evidence for negative interaction effects even among noncompeting advertisements. Calder and Sternthal (1980) show that when subjects are repeatedly exposed to an ad embedded in a collection of commercials over a period of time, they have negative evaluations of the ad and the associated product. Attention to a specific message theme is theorized to diminish if the message is dominated by a clutter of other message themes. Moreover, there is evidence that the persuasive impact of advertising is much less reliable when a target ad is presented in an environment that includes other ads (Belch 1982, Burke and Srull 1988, Ray and Sawyer 1971, Rethans et al. 1986, Malaviya et al. 1999). It is interesting to see that both price ads and product offer ads with rational appeal do not have significant interaction effects. This may be suggesting that attention to rational price offer or product offer messages is less diminished by message clutter. Emotional ads appear to have negative interaction effects on goodwill even though they reduce wearout.

In Table 7, we present the estimates of our model without interaction effects. We find that this model is rejected by log Bayes factor criterion. The copy wearout and repetition wearout effects are consistent with those in the earlier model. We find that rational ads do have faster copy wearout but slower repetition wearin. The forgetting parameter is 0.0784, significant and higher than that obtained earlier. In fact, without the interaction terms all the wearout effects are larger in magnitude, thus suggesting bias if one does not capture these interaction effects. The magnitude of the bias averages 58% with a range from 21% to 146%.

Table 7 Estimates by Message Type (No Interactions)

Parameters	Estimate	Standard deviation	95% HPDI	
Price per minute	−23.8591	15.290	−56.999	−3.1266
Competitive advertising	−0.1664	0.2391	−0.5630	0.1774
Capacity (number of lines)	2.1309	1.3991	0.1130	5.0606
Forgetting rate δ	0.0784	0.0345	0.0419	0.1301
Initial goodwill G_0	16.5864	2.6498	14.219	23.106
Observation variance σ_ε	0.3483	0.6733	0.0962	1.8670
System variance (goodwill) $\sigma_{\varepsilon 1}$	0.0815	0.0107	0.0658	0.1005
Call stimulation				
Copy wearout c	0.5119	0.1297	0.3446	0.6758
Repetition wearout w	−0.0811	0.0282	−0.1160	−0.0441
Initial quality q_{10}	0.1330	0.0635	0.0374	0.2475
System variance (ad quality) $\sigma_{\varepsilon 2}$	0.0020	2.5920e−4	0.0016	0.0024
Price offer				
Copy wearout c	0.7097	0.1775	0.4388	0.9267
Repetition wearout w	−0.1349	0.0357	−0.1795	−0.0843
Initial quality q_{20}	0.1064	0.0639	0.0379	0.2601
System variance (ad quality) $\sigma_{\varepsilon 3}$	1.9891e−4	2.5992e−5	1.6072e−4	2.4427e−4
Product offer				
Copy wearout c	0.6196	0.1225	0.4663	0.7762
Repetition wearout w	−0.1010	0.0252	−0.1322	−0.0695
Initial quality q_{30}	0.1990	0.0929	0.0968	0.4218
System variance (ad quality) $\sigma_{\varepsilon 4}$	7.9434e−4	1.0497e−4	6.3979e−4	9.809e−4
Reconnection				
Copy wearout c	0.2749	0.0586	0.1850	0.3344
Repetition wearout w	−0.0601	0.0130	−0.0738	−0.0402
Initial quality q_{40}	0.2222	0.1069	0.1119	0.4907
System variance (ad quality) $\sigma_{\varepsilon 5}$	8.0700e−5	1.0522e−5	6.4914e−5	9.499e−5
Reassurance				
Copy wearout c	0.2807	0.0559	0.2028	0.3424
Repetition wearout w	−0.0711	0.0155	−0.0893	−0.0494
Initial quality q_{50}	0.2203	0.1034	0.1020	0.4722
System variance (ad quality) $\sigma_{\varepsilon 6}$	1.5989e−4	2.1034e−5	1.2847e−4	1.9690e−4

In Table 8, we present estimates of our model using aggregate advertising GRP instead of considering the GRP of the different themes separately. Using log Bayes factor (87.712), we can reject this model in favor of our proposed model. We find that the effect of price, competitive advertising, and capacity are consistent with earlier results. In the aggregate model, we find that the forgetting rate is 0.93, which is very high compared to estimates obtained in the above models as well as the estimates obtained in earlier studies. Note that the forgetting parameter is identified by periods when there is no advertising. When we add up the GRPs of all themes of ads (Figure 2), there are very few weeks in which there is no advertising. Thus, the aggregate model will provide a biased estimate. We find the copy wearout parameter is negative and the repetition wearout parameter is positive. These results suggest that when there are multiple themes being advertised, an aggregate model might provide misleading interpretation.

Table 8 Estimates of Aggregate Model

Parameters	Estimate	Standard deviation	95% HPDI	
Price per minute	−10.4908	4.3619	−17.823	−3.2900
Competitive advertising	−0.3816	0.1523	−0.6349	−0.1316
Capacity (number of lines)	1.5208	0.3979	0.8602	2.1856
Forgetting rate δ	0.9322	0.0523	0.8594	0.9911
Initial goodwill G_0	1.5253	0.3622	0.9251	2.1201
Observation variance σ_ε	0.8947	0.1658	0.6600	1.1813
System variance (goodwill) $\sigma_{\varepsilon 1}$	0.0293	0.0039	0.0234	0.0362
Aggregated				
Copy wearout c	−0.0293	0.0164	−0.0494	−0.0093
Repetition wearout w	2.8089e−4	6.4801e−5	1.9014e−4	3.7902e−4
Initial quality q_{10}	0.2279	0.0547	0.1368	0.3162
System variance (ad quality) $\sigma_{\varepsilon 2}$	7.4072e−4	9.6713e−5	5.959e−4	9.1178e−4

Figure 3 shows the evolution of the goodwill parameter (G_t) and the five quality parameters (q_{it}).

To provide a measure of the model's predictive performance, we use the parameters from our proposed model in Table 4 and predict the number of hours of talk time. The plot of predicted versus actual hours of talk time is given in Figure 4. The plot indicates a very good fit of the model with the data across 114 weeks. The predicted values track the variations in actual data quite closely.

5.1. Advertising Policy Implications

Given our estimates in Table 4, we can reconsider the advertiser's budget allocation decisions across the five advertising themes. For example, we estimated that over the 114-week period of our sample, price and product ads had the two highest copy wearout estimates but the lowest repetition wearouts. On the other hand, reconnection and reassurance advertising themes had the lowest copy wearouts and call stimulation and reassurance had the highest repetition wearouts. What are the implications of these results on budget allocation in each of the 114 weeks? More importantly, could the advertising have been more efficient in generating greater demand for telephone services?

To answer the above questions, we develop a model to reallocate the total advertising GRPs (b_t) in each period across the five ($m = 5$) ad themes to maximize total expected telephone service demand over 114 weeks. We solve the following large-scale nonlinear optimization problem P1:

$$\max_{A_{11}\cdots A_{1m}\cdots A_{T1}\cdots A_{Tm}} \sum_{t=1}^{T} E(y_t \mid D_{t-1})$$

$$\text{s.t.} \quad \sum_{i=1}^{m} A_{it} \le b_t, \quad A_{it} \ge 0 \quad t = 1, \ldots, T$$

$$\operatorname{Var}(y_t \mid D_{t-1}) \le \sigma_t^2 \quad t = 1, \ldots, T$$

where $E(y_t \mid D_{t-1})$ and $\operatorname{Var}(y_t \mid D_{t-1})$ are the expectation and variance of the one-step-ahead forecast distribution (see Appendix A, Equation A3) and A_{it} the GRP for an ad theme at time t. Thus, we solve a program that involves 570 (5×114) variables and 114 (114 weeks) linear constraints.

The problem (P1) is solved in SNOPT, a set of Fortran routines developed by Stanford Optimization Laboratory. These routines are called within MATLAB. SNOPT is a general purpose system for solving optimization problems involving many variables and constraints, and so is particularly suited for our problem. It uses a sequential quadratic programming algorithm that obtains search directions from a sequence of quadratic programming subproblems (Gill et al. 2002). Given the size of the problem it solves, SNOPT finds solutions that are locally optimal and, ideally, any nonlinear objective functions should be smooth. We use multiple start values and choose solutions that give allocations that represent an improvement in advertising efficiency. We caution the readers that the solution may not be optimal but provides a better allocation.

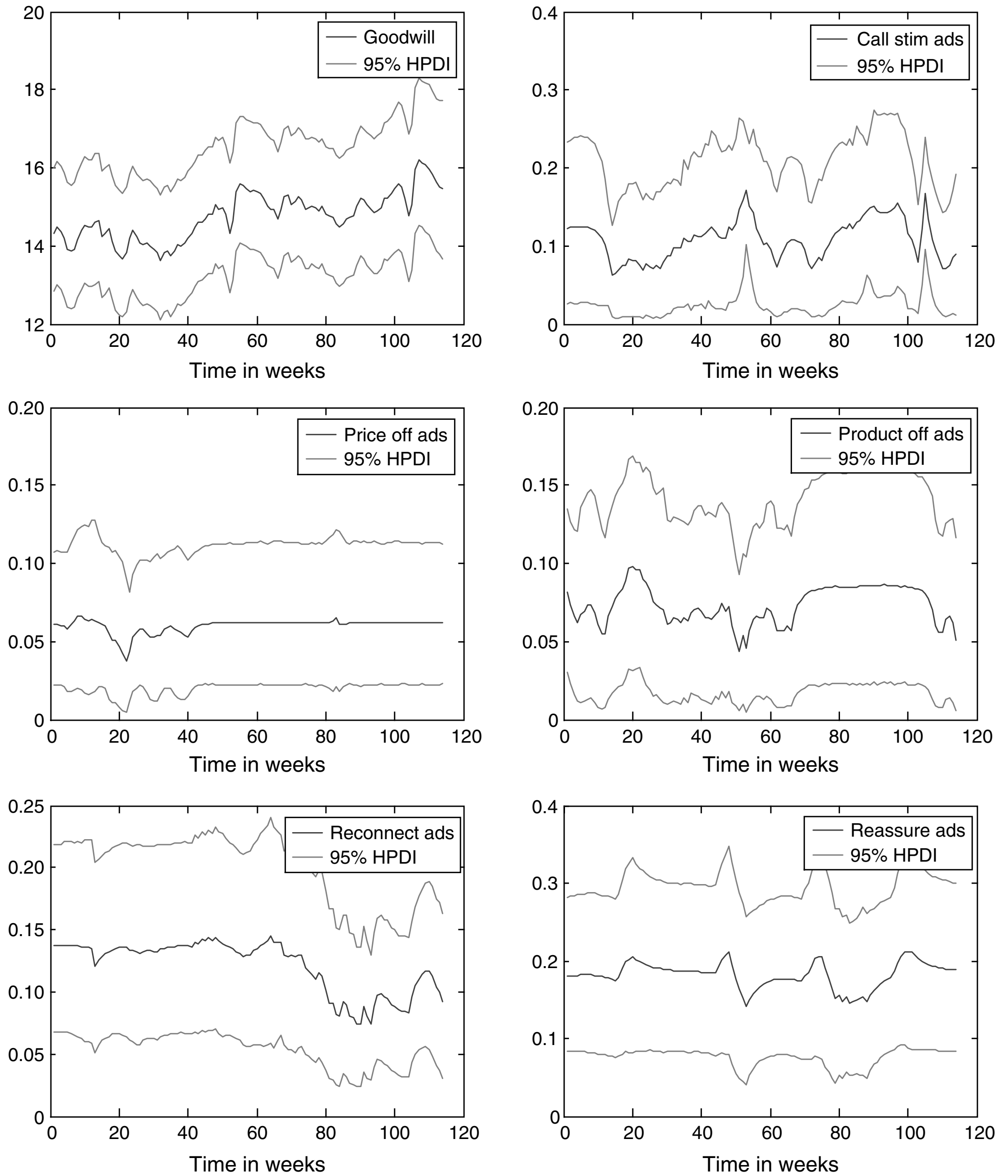
One such allocation is reported in Table 9 and depicted in Figure 5. The horizontal bar graph compares the "improved" and the actual allocation of resources (measured in total GRPs) across the different themes. Our model suggests increasing advertising expenditure on reconnect and reassurance ads while decreasing on the other three themes. The percentage changes in allocation of GRPs are quite large as seen in Table 9. They range from a low of 50% to a high of 276% change.

A reallocation of advertising along the suggested lines would generate an additional 35.82 million hours of calling time which represents a 2% increase over the current level of demand. We thus demonstrate how our model can be used to reallocate resources over different themes to improve the desired outcome, whether it is awareness as in Naik et al. (1998) or demand as in our case.

6. Conclusion

We have developed a model of demand that considers the dynamic effects of multiple themes of television advertising. We believe that this is a first attempt to help managers allocate their advertising resources across different themes of advertising. The model considers wearout effects—both copy wearout and repetition wearout—as well as forgetting. Advertising is assumed to affect goodwill which in turn affects demand for a product or a service. A modification of the Nerlove-Arrow specification is employed because it has been shown to perform the best (Naik et al. 1998). Our model extends the specification in Naik et al. (1998) by considering the evolution of goodwill and advertising effectiveness over time for

Figure 3 Plot of the Dynamic Effects of the Coefficients with 95% HPDI

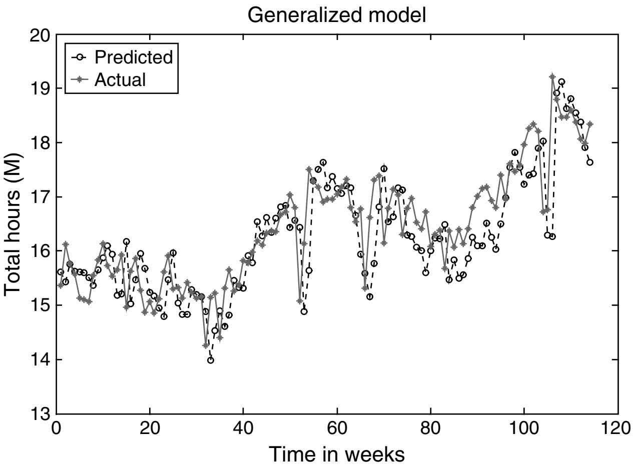


different ad themes. Further, we employ dynamic linear Bayesian estimation techniques (West and Harrison 1997), a relatively new method, to estimate our model parameters. We also develop a model to show how reallocation of resources can be done using the parameters of our proposed model.

We find that modeling different advertising efforts on different themes yields better insights and unbiased estimates relative to a model that aggregates the advertising effort. We find that copy wearout effects of different themes are positive and systematically different. Based on our ad hoc classification of ad themes as either emotional or rational ads, we find some support for the premise proposed in behavioral literature (Pechmann and Stewart 1990) that emotional ads tend to wear out faster than nonemotional ads. Repetition wearout parameters are negative, suggesting wearin effects for this data. We find some evidence that repetition wearout is lower if the number of copy executions is higher.

Thus, changing the execution of the message appears to refresh the message and mitigates wearout. There is evidence of positive forgetting effects when there is no advertising. Thus, our model is consistent with earlier findings on the effects of advertising with repetition. The main value of our model comes from the fact that we have linked demand to the wearout and forgetting effects and so have provided a way for managers to use these parameters and insights in making allocation decisions. Our policy experiment suggests that a reallocation of advertising effort across the different themes could result in a higher payoff.

Figure 4 Plot of Predicted and Actual Hours of Talk Time Over 114 Weeks



There are a few limitations of our model. In the proposed model, we assume that advertising is an exogenous variable. Since managers allocate advertising in a strategic manner, one can make a case that it should be treated as an endogenous variable as we have done for price. This control for the endogeneity of advertising can be modeled if good instruments were available in the data. We have not considered the dynamic effects of competitive advertising, which could be modeled analogous to the different themes. Given data on competitors' demand, this extension can be attempted in future research. In our data, the firm is a monopolist in its business and so we have not modeled competition explicitly. In other product categories, this would be an area for future research.

Our model can be extended to account for changes in copy and a model can be developed for allocation of resources across different copy. This is a challenging problem since the number of executions even for a given theme can be large and one needs

Figure 5 Actual vs. Model-Based Allocation of GRPs for Each Advertising Theme

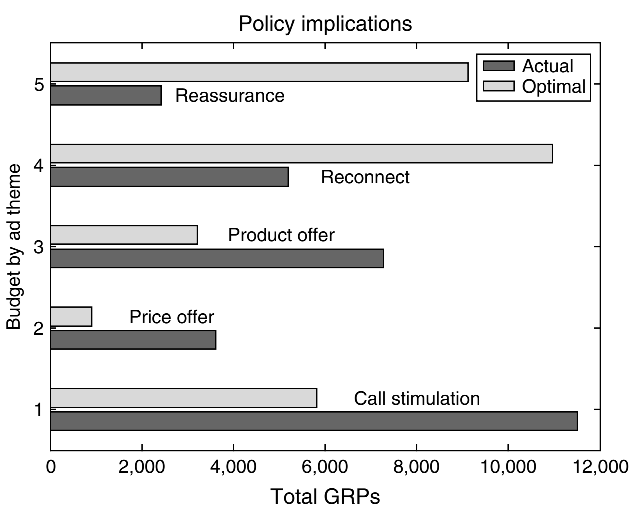


Table 9 Comparison of Actual and Model-Based GRPs

Theme of advertising	Actual GRPs	Model-based GRPs	% Change (%)
Call stimulation	11,509	5,815	−49.5
Price offer	3,611	914	−74.7
Product offer	7,281	3,218	−55.8
Reconnect	5,201	10,962	+110.7
Reassurance	2,422	9,115	+276.3

to also find a way to incorporate content of advertising into the model. Other avenues for future research are to extend our model to develop a media allocation model and to also consider the interaction between different themes and different media. These are important issues and our paper presents a possible starting point to address them.

Acknowledgment

The authors acknowledge first author Frank M. Bass, who died December 1, 2006, for his lifetime of contributions to marketing science.

Appendix A

This appendix provides an overview of the posterior sampling algorithm. The sampling scheme shown in §A.1 is direct application of DLM theory (West and Harrison 1997, Chapter 4) and Gibbs sampling for state space models developed by Carter and Kohn (1994) and Fruhwirth-Schnatter (1994). Our model belongs to the class of linear state space models in which components of the transition matrix ($\mathbf{H}_t$) and the variance components ($\sigma_\varepsilon^2, \mathbf{W}$) are treated as parameters to be estimated along with the sequence of state vectors ($\mathbf{\Phi}_t$) over time. We begin the forward filtering step with the most recent values of $\mathbf{m}_0$, $\mathbf{C}_o$, $\mathbf{W}$, σ_ε^2, $\mathbf{\Lambda}$, and $\boldsymbol{\beta}$:

A.1. Sampling from $p(\mathbf{\Phi}_t \mid \mathbf{D}_t)$

Forward Filtering. We use the standard DLM framework (Equations A1 through A5) to infer the posterior distribution $\mathbf{\Phi}_t \mid \mathbf{D}_t$ over time, where $\mathbf{D}_t = \{y_t, \mathbf{D}_{t-1}\}$ includes all information available to the researcher at time t. The posterior $\mathbf{\Phi}_t \mid \mathbf{D}_t$ is then determined using standard mutivariate normal theory:

Posterior distribution for $\mathbf{\Phi}_{t-1}$,

$$\mathbf{\Phi}_{t-1} \mid \mathbf{D}_{t-1} \sim \mathbf{N}(\mathbf{m}_{t-1}, \mathbf{C}_{t-1}). \tag{A1}$$

Prior distribution for $\mathbf{\Phi}_t$: $\mathbf{\Phi}_t \mid \mathbf{D}_{t-1} \sim \mathbf{N}(\mathbf{a}_t, \mathbf{R}_t)$, where

$$\mathbf{a}_t = \mathbf{H}_t \mathbf{m}_{t-1} + \mathbf{u}_t, \quad \text{and} \quad \mathbf{R}_t = \mathbf{H}_t' \mathbf{C}_{t-1} \mathbf{H}_t + \mathbf{W}. \tag{A2}$$

Prior one-step-ahead forecast distribution: $y_t \mid \mathbf{D}_{t-1} \sim N(f_t, P_t)$,

$$f_t = \mathbf{F}_t \mathbf{a}_t + \boldsymbol{\beta} \mathbf{X}_t, \qquad P_t = \mathbf{F}_t' \mathbf{R}_t \mathbf{F}_t + \sigma_\varepsilon^2. \tag{A3}$$

$$\Phi_t, y_t \mid \mathbf{D}_{t-1} \sim N\left[\begin{Bmatrix} \mathbf{a}_t \\ f_t \end{Bmatrix}, \begin{Bmatrix} \mathbf{R}_t & \mathbf{R}_t\mathbf{F}_t \\ (\mathbf{R}_t\mathbf{F}_t)' & P_t \end{Bmatrix}\right], \quad \text{with}$$
$$\begin{aligned} \mathrm{Cov}(\Phi_t, y_t) &= \mathrm{Cov}(\mathbf{\Phi}_t, \mathbf{F}_t\mathbf{\Phi}_t + \boldsymbol{\beta}\mathbf{X}_t + \varepsilon_t) \\ &= \mathrm{Var}(\mathbf{\Phi}_t \mid \mathbf{D}_{t-1})\mathbf{F}_t = \mathbf{R}_t\mathbf{F}_t. \end{aligned} \tag{A4}$$

Posterior distribution for $\mathbf{\Phi}_t$: $\mathbf{\Phi}_t \mid y_t, \mathbf{D}_t \sim N(\mathbf{m}_t, \mathbf{C}_t)$, making use of the marginal properties of the normal in (A4),

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t(y_t - f_t), \quad \mathbf{A}_t = \mathbf{R}_t\mathbf{F}_t\mathbf{P}_t^{-1}, \quad \text{where } \mathbf{C}_t = \mathbf{R}_t(\mathbf{I} - \mathbf{F}_t\mathbf{A}_t). \quad \text{(A5)}$$

Backward Smoothing. We derive the backward smoothing algorithm by using (A1) through (A2) to write down the joint distribution of the parameters at t and $t-1$, given information $\mathbf{D}_{t-1}$ to obtain Equation (A6):

$$\mathbf{\Phi}_t, \mathbf{\Phi}_{t-1} \mid \mathbf{D}_{t-1} \sim \mathbf{N}\left[\begin{Bmatrix} \mathbf{a}_t \\ \mathbf{m}_{t-1} \end{Bmatrix}, \begin{Bmatrix} \mathbf{R}_t & \mathbf{G}_t\mathbf{C}_{t-1} \\ (\mathbf{G}_t\mathbf{C}_{t-1})' & \mathbf{C}_{t-1} \end{Bmatrix}\right]. \quad \text{(A6)}$$

Then, using standard multivariate normal theory, we obtain the conditionals:

$$\mathbf{\Phi}_{t-1} \mid \mathbf{\Phi}_t, \mathbf{D}_{t-1} \sim \mathbf{N}(\mathbf{m}_{t-1} + \mathbf{B}_t(\mathbf{\Phi}_t - \mathbf{a}_t), \mathbf{C}_{t-1} - \mathbf{B}_t'\mathbf{R}_t\mathbf{B}_t), \quad \text{(A7)}$$

where $\mathbf{B}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{R}_t^{-1}$ and $\mathbf{\Phi}_t$ is a random draw from the posterior $\mathbf{N}(\mathbf{\Phi}_t \mid \mathbf{D}_t)$.

We can now do a retrospective analysis using results derived in Equations (A6) and (A7). That is, we take expectations and variance over all possible draws of the posterior of $\mathbf{\Phi}_t$ given information at time t, $\mathbf{D}_t$. In other words, at any time (say, $t-1$) in the analysis the researcher can update his posterior belief about the effects of advertising spending given his new information at time t. The result is our backward smoothing algorithm:

$$\begin{aligned} \mathbf{E}[\mathbf{\Phi}_{t-1} \mid \mathbf{D}_t] &= \mathbf{E}[\mathbf{E}[\mathbf{\Phi}_{t-1} \mid \mathbf{\Phi}_t, \mathbf{D}_{t-1}] \mid \mathbf{D}_t] \\ &= \mathbf{m}_{t-1} + \mathbf{B}_t(\mathbf{m}_t - \mathbf{a}_t) \end{aligned} \quad \text{(A8)}$$

and

$$\begin{aligned} \mathrm{Var}(\mathbf{\Phi}_{t-1} \mid \mathbf{D}_t) &= \mathbf{E}[\mathrm{Var}[\mathbf{\Phi}_{t-1} \mid \mathbf{\Phi}_t, \mathbf{D}_{t-1}] \mid \mathbf{D}_t] \\ &\quad + \mathrm{Var}[\mathbf{E}[\mathbf{\Phi}_{t-1} \mid \mathbf{\Phi}_t, \mathbf{D}_{t-1}] \mid \mathbf{D}_t] \\ &= \mathbf{C}_{t-1} - \mathbf{B}_t(\mathbf{R}_t - \mathbf{C}_t)\mathbf{B}_t'. \end{aligned} \quad \text{(A9)}$$

We use the above (Equations A1 through A9) to sample the state vector of goodwill and advertising copy quality as follows:

Simulation for $\mathbf{\Phi}^T$.

Step 1. For $t = 1, \ldots, T$, compute the moments $(\mathbf{m}_t, \mathbf{C}_t)$ for the multivariate normal $p(\mathbf{\Phi}_t \mid \mathbf{D}_t, \sigma_\varepsilon^2, \mathbf{W})$ by applying the sequential updating procedure described in the above forward filter section (Equations A2, A3, and A5).

Step 2. At the end of the series $(t = T)$, sample $\mathbf{\Phi}_T$ from the posterior distribution: $p(\mathbf{\Phi}_T \mid \mathbf{D}_T, \sigma_\varepsilon^2, W) = \mathbf{N}(\mathbf{\Phi}_T \mid \mathbf{m}_T, \mathbf{C}_T)$.

Step 3. For $t = T, \ldots, 1$, sample $p(\mathbf{\Phi}_{t-1} \mid \mathbf{\Phi}_t, \sigma_\varepsilon^2, W)$ conditional on the latest draw $\mathbf{\Phi}_t$.

The results are the draws $\mathbf{\Phi}^T = \{\mathbf{\Phi}_0, \mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_T\}$ from the full conditional posterior.

A.2. Sampling from $p(\mathbf{W}, \sigma_\varepsilon^2, \mathbf{\Lambda}, \boldsymbol{\beta} \mid \mathbf{\Phi}^T, \mathbf{y}^T)$

Conditional on all the states and the data $\{\mathbf{\Phi}^T, \mathbf{y}^T\}$, our DLM Equations (9) and (10) simplify to a linear multivariate system with unknowns, parameters $\{\mathbf{\Lambda}, \boldsymbol{\beta}\}$, and variance components $\{\mathbf{W}, \sigma_\varepsilon^2\}$. Consequently, the Gibbs sampler step to estimate the joint posterior of the nonstate parameters conditional on the data and all states $p(\mathbf{W}, \sigma_\varepsilon^2, \mathbf{\Lambda}, \boldsymbol{\beta} \mid \mathbf{\Phi}^T, \mathbf{y}^T)$ is very straightforward. We refer the reader to Gelfand and Smith (1990) for an overview. For the interested reader, we note several DLM features that highlight the simplicity of this step:

(1) The observation Equation (9) and system Equation (10) errors are mutually independent (see West and Harrison 1997). Thus, conditional on $(\mathbf{\Lambda}, \boldsymbol{\beta})$, we can sample $\{\mathbf{W}, \sigma_\varepsilon^2\}$ independently.

(2) Similarly, conditional on $\{\mathbf{\Phi}^T, \mathbf{y}^T\}$, the states and the data $\{\mathbf{\Delta}, \boldsymbol{\beta}\}$ are independent and are thus sampled separately.

Appendix B

To show identification in our DLM, we need to investigate whether enough prior structure exists in the system of Equations (11) and (12). First, we need to construct an observationally equivalent system $M_1 = \{\mathbf{F}_{1t}, \mathbf{H}_{1t}, \sigma_\varepsilon^2, \mathbf{W}_1\}$, where observability (Crassidis and Junkins 2004) relates to our ability to recover all state vectors $\mathbf{\Phi}^T = \{\mathbf{\Phi}_0, \mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_T\}$, given observations $\mathbf{y}^T = \{y_1, y_2, \ldots, y_T\}$. An equivalent system is achieved by substituting the linear transformation $\mathbf{\Phi}_{1t} = \mathbf{L}\mathbf{\Phi}_t$ into Equations (11) and (12), where $\mathbf{L}_t$ is a $(m+1) \times (m+1)$ nonsingular matrix (West and Harrison 1997, Chapter 4). Thus, we have:

$$\mathbf{F}_{1t} = \mathbf{F}_t\mathbf{L}^{-1} \quad \text{(B1)}$$

$$\mathbf{H}_{1t} = \mathbf{L}\mathbf{H}_t\mathbf{L}^{-1} \quad \text{(B2)}$$

$$\mathbf{W}_1 = \mathbf{L}\mathbf{W}\mathbf{L}' \quad \text{(B3)}$$

$$\mathbf{u}_{1t} = \mathbf{L}\mathbf{u}_t. \quad \text{(B4)}$$

We illustrate identification for a model with a single message theme (i.e., $m = 1$; Naik et al. 1998). We will show that enough prior structure exists in $\mathbf{F}_t$ and $\mathbf{H}_t$ such that all transformations are precluded except $\mathbf{L} = \mathbf{I}$. The results generalize to the multitheme case.

Define the nonsingular matrix

$$\mathbf{L} = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix}$$

and recall and note constraints in $\mathbf{F}_t$ Equations (11).

From Equation (B1) we have

$$\begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix}.$$

Thus, $l_{11} = 1$, $l_{12} = 0$, and $l_{22} \neq 0$ ($\mathbf{L}$ is nonsingular).

Now, recall and note the constraints in the system matrix $(\mathbf{H}_t)$ from Equation (12), when advertising is off $I(A) = 0$. From Equation (B2), we have:

$$\underbrace{\begin{pmatrix} h_{11} & 0 \\ 0 & h_{22} \end{pmatrix}}_{\mathbf{H}_{1t}}\begin{pmatrix} 1 & 0 \\ l_{21} & l_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ l_{21} & l_{22} \end{pmatrix}\underbrace{\begin{pmatrix} 1-\delta & 0 \\ 0 & 1-c-\delta \end{pmatrix}}_{\mathbf{H}_t}.$$

By inspection, $\mathbf{H}_{1t} = \mathbf{H}_t$ and $l_{21} = 0$.

Given our specification of the quality evolution (5) and $\mathbf{H}_{1t} = \mathbf{H}_t$, we must have $\mathbf{u}_{1t} = \mathbf{L}\mathbf{u}_t$ and as a result, $\mathbf{l}_{22} = 1$.

That is, enough prior structure exists in $\mathbf{F}_t$ and $\mathbf{G}_t$ such that all transformations are precluded except $\mathbf{L} = \mathbf{I}$. QED

Note that Naik et al. (1998) is a model of pulsing. We now consider identification if this model was applied to situations in which there is no pulsing $I(A) = 1$:

$$\mathbf{H}_{1t} = \begin{pmatrix} 1 & 0 \\ l_{21} & l_{22} \end{pmatrix} \underbrace{\begin{pmatrix} 1-\delta & g(A) \\ 0 & 1-a(A) \end{pmatrix}}_{\mathbf{H}_t} \begin{pmatrix} 1 & 0 \\ l_{21} & l_{22} \end{pmatrix}^{-1}$$

$$= \frac{1}{l_{22}} \begin{pmatrix} al_{22} - bl_{21} & b \\ al_{21}l_{22} - bl_{21}^2 - cl_{21}l_{22} & bl_{21} + cl_{22} \end{pmatrix},$$

where $a = 1-\delta$, $b = g(A)$, and $c = 1 - a(A)$.

Given the structure of $\mathbf{H}_t$, $al_{21}l_{22} - bl_{21}^2 - cl_{21}l_{22} = 0$. Thus, $l_{21} = 0$ or $l_{21} = ((a-c)/b)l_{22}$. Note that if we take $l_{21} = ((a-c)/b)l_{22}$, $\mathbf{H}_{1t}(1,1) = c$ and $\mathbf{H}_{1t}(2,2) = a$, which is inconsistent with evolution of quality or goodwill (Equations 4 and 5). Thus, we take $l_{21} = 0$. If there is no pulsing, Equation (B4) provides no information since $\mathbf{u}_{1t} = \mathbf{u}_t = \mathbf{0}$. Consider Equation (B3):

$$\mathbf{W}_1 = \begin{pmatrix} 1 & 0 \\ 0 & l_{22} \end{pmatrix} \underbrace{\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}}_{\mathbf{W}} \begin{pmatrix} 1 & 0 \\ 0 & l_{22} \end{pmatrix}'$$

$$= \begin{pmatrix} w_{11} & l_{22}w_{12} \\ l_{22}w_{21} & l_{22}^2 w_{22} \end{pmatrix}.$$

Identification can be achieved if we restrict the system variance diagonal to be equal, giving $l_{22} = 1$ or $l_{22} = -1$. We reject $l_{22} = -1$ since this implies $\mathbf{H}_{1t}(1,1) < 0$, which is impossible given the constraint in $\mathbf{H}_t(1,1) > 0$. Thus, we have $\mathbf{L} = \mathbf{I}$. QED

Note that imposing the above constraint leads to model identification.

References

Aaker, D. A., J. M. Carman. 1982. Are you overadvertising? *J. Advertising Res.* **22**(4) 57–70.

Allenby, G., P. Rossi. 2003. Bayesian statistics and marketing. *Marketing Sci.* **22**(3) 304–328.

Axelrod, J. 1980. Advertising wearout. *J. Advertising Res.* **20** 65–74.

Bass, F. M., D. G. Clarke. 1972. Testing distributed lag models of advertising effects. *J. Marketing Res.* **9** 298–308.

Bass, F. M., R. P. Leone. 1983. Temporal aggregation, the data interval bias, and empirical estimation of bimonthly relations from annual data. *Management Sci.* **29** 1–11.

Belch, G. E. 1982. The effects of television commercial repetition on cognitive response and message acceptance. *J. Consumer Res.* **9** 56–96.

Berlyne, D. E. 1970. Novelty, complexity, and hedonic value. *Perception Psychophys.* **8** 279–286.

Blackburn, J. D., K. J. Clancy. 1982. LITMUS: A new product planning model. A. A. Zoltners, ed. *Marketing Planning Models*. North Holland, New York, 43–62.

Blair, M. H., M. J. Rabuck. 1998. Advertising wearin and wearout: Ten years later—More empirical evidence and successful practice. *J. Advertising Res.* **38**(5) 7–18.

Blattberg, R., J. Golanty. 1978. TRACKER: An early test market forecasting and diagnostic model for new product planning. *J. Marketing Res.* **15** 192–202.

Blattberg, R. C., A. P. Jeuland. 1981. A micro-modeling approach to investigate the advertising-sales relationship. *Management Sci.* **27**(9) 988–1004.

Broadbent, S. 1984. Modeling with ad stock. *J. Market Res. Soc.* **16** 295–312.

Bronnenberg, B. J. 1998. Advertising frequency decisions in a discrete Markov process under a budget constraint. *J. Marketing Res.* **35**(August) 399–406.

Bronnenberg, B. J., V. Mahajan, W. Vanhonacker. 2000. The emergence of market structure in new repeat-purchase categories: The interplay of market share and retailer distribution. *J. Marketing Res.* **37** 16–31.

Burke, R., T. K. Srull. 1988. Competitive inference and consumer memory for advertising. *J. Consumer Res.* **15** 55–67.

Calantone, R., A. G. Sawyer. 1978. The stability of benefit segments. *J. Marketing Res.* **15**(August) 395–404.

Calder, B., B. Sternthal. 1980. Television commercial wearout: An information processing view. *J. Marketing Res.* **17** 173–186.

Carroll, J. D., P. G. Green, W. S. DeSarbo. 1979. Optimising the allocation of a fixed resource: A single model and its experimental test. *J. Marketing* **43**(January) 51–57.

Carter, C., R. Kohn. 1994. On Gibbs sampling for state space models. *Biometrika* **81** 541–553.

Clarke, D. G. 1976. Econometric measurement of the duration of advertising effect on sales. *J. Marketing Res.* **13**(November) 345–357.

Congdon, P. 2001. *Bayesian Statistical Modelling*. John Wiley and Sons, Chichester, UK.

Corkindale, D., J. Newall. 1978. *Advertising Threshold and Wearout*. MCB Publications, Bradford, UK.

Crassidis, J. L., J. L. Junkins. 2004. *Optimal Estimation of Dynamic Systems*. Chapman and Hall, New York.

Doyle, P., J. Saunders. 1990. Multiproduct advertising budgeting. *Marketing Sci.* **9**(2) 97–113.

Dubé, J.-P., P. Manchanda. 2005. Differences in dynamic brand competition across markets: An empirical analysis. *Marketing Sci.* **24**(1) 81–95.

Eastlack, J. O., A. G. Rao. 1989. Advertising experiments at the Campbell Soup Company. *Marketing Sci.* **8**(1) 57–71.

Eskin, G. J., P. H. Baron. 1977. Effects of price and advertising in test–Market experiments. *J. Marketing Res.* **14**(November) 499–508.

Feinberg, F. 1992. Pulsing policies for aggregate advertising models. *Marketing Sci.* **11**(3) 221–234.

Fruhwirth-Schnatter, S. 1994. Data augmentation and dynamic linear models. *Time Ser. Anal.* **15** 183–202.

Gelfand, A. E., A. F. M. Smith. 1990. Sampling-based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.* **85** 972–985.

Gill, P. E., W. Murray, M. Saunders. 2002. *User's Guide for SNOPT Version 7, Software for Large Scale Nonlinear Programming*, http://www.sbsi-sol-optimize.com/manuals/SNOPT Manual.pdf.

Grass, R., W. H. Wallace. 1969. Satiation effect of TV commercials. *J. Advertising Res.* **1** 1–13.

Greenberg, A., C. Suttoni. 1973. TV commercial wearout. *J. Advertising Res.* **13** 47–54.

Greyser, S. A. 1973. Irritation in advertising. *J. Advertising Res.* **13**(February) 3–10.

Hanssens, D. M., L. J. Parsons, R. L. Schultz. 1990. *Market Response Models: Econometric and Time Series Analysis*. Kluwer, Boston, MA.

Harvey, A. 1994. *Forecasting Structural Time Series and the Kalman Filter.* Cambridge University Press, Cambridge, UK.

Hitchon, J., E. Thorson, X. Zhao. 1988. Advertising repetition as a component of the viewing environment: Impact of emotional executions on commercial reception. Working paper, School of Journalism and Mass Communication, University of Wisconsin at Madison, Madison, WI.

Jagpal, H. S., E. F. Sudit, H. D. Vinod. 1979. A model of sales response to advertising interactions. *J. Advertising Res.* **19**(3) 41–47.

Jedidi, K., C. F. Mela, S. Gupta. 1999. Managing advertising and promotion for long-run profitability. *Marketing Sci.* **18**(1) 1–22.

Krishnamurthi, L., S. P. Raj. 1985. The effect of advertising on consumer price sensitivity. *J. Marketing Res.* **22**(May) 119–129.

Lambin, J. J. 1969. Measuring the profitability of advertising: An empirical study. *J. Indust. Econom.* **17**(April) 86–103.

Lancaster, T. 2004. *An Introduction to Modern Bayesian Econometrics*. Blackwell Publishing, Oxford, UK.

Leichty, J. C., D. K. H. Fong, W. S. DeSarbo. 2005. Dynamic models incorporating individual heterogeneity: Utility evolution in conjoint analysis. *Marketing Sci.* **24**(2) 285–293.

Little, J. D. C. 1975. BRANDAID: A marketing mix model, Part 1—Structure. *Oper. Res.* **23** 628–655.

Little, J. D. C. 1979. Aggregate advertising model: The state of the art. *Oper. Res.* **27**(4) 629–667.

Lodish, L. M., M. Abraham, S. Kalmenson, J. Livelsberger, B. Lubetkin, B. Richardson, M. E. Stevens. 1995. How TV advertising works: A meta-analysis of 389 real world split cable TV advertising experiments. *J. Marketing Res.* **32**(May) 125–139.

MacInnis, D., L. L. Price. 1987. The role of imagery in information processing: Review and extensions. *J. Consumer Res.* **13**(March) 473–491.

MacInnis, D. J., A. G. Rao, A. M. Weiss. 2002. Assessing when increased media weight of real-world advertisements helps sales. *J. Marketing Res.* **39**(November) 391–407.

Mahajan, V., E. Muller. 1986. Advertising pulsing policies for generating awareness of new products. *Marketing Sci.* **5**(2) 86–106.

Mahajan, V., E. Muller, S. Sharma. 1984. An empirical comparison of awareness forecasting models for new product introduction. *Marketing Sci.* **3** 179–206.

Malaviya, P., J. Meyers-Levy, B. Sternthal. 1999. Ad repetition in a cluttered environment: The influence of type of processing. *Psych. Marketing* **16**(2) 99–118.

Mela, C. F., S. Gupta, D. R. Lehmann. 1997. The long term impact of promotions and advertising on brand choice. *J. Marketing Res.* **34**(May) 248–261.

Montgomery, D. B., M. C. Moore, J. E. Urbany. 2005. Reasoning about competitive reactions: Evidence from executives. *Marketing Sci.* **24**(1) 138–149.

Naik, P. A., M. K. Mantrala, A. G. Sawyer. 1998. Planning media schedules in the presence of dynamic advertising quality. *Marketing Sci.* **17**(3) 214–235.

Naik, P., K. Raman, R. S. Winer. 2005. Planning marketing-mix strategies in the presence of interaction effects. *Marketing Sci.* **24**(1) 25–34.

Neelamegham, R., P. Chintagunta. 2004. Modeling and forecasting the sales of technology products. *Quant. Marketing Econom.* **2**(3) 195–232.

Nerlove, M., K. Arrow. 1962. Optimal advertising policy under dynamic conditions. *Economica* **29**(May) 129–142.

Palda, K. S. 1964. *The Measurement of Cumulative Advertising Effects*. Prentice Hall, Englewood Cliffs, NJ.

Pechmann, C., D. W. Stewart. 1990. Advertising repetition: A critical review of wearin and wearout. MSI report.

Rao, A. G., P. B. Miller. 1975. Advertising/sales response functions. *J. Advertising Res.* **15**(April) 7–15.

Ray, M. L., A. G. Sawyer. 1971a. Behavioral measurement for marketing models: Estimating the effects of advertising repetition for media planning. *Management Sci.* **18**(4, Part 2) 73–89.

Ray, M. L., A. G. Sawyer. 1971b. Repetition in media models: A laboratory technique. *J. Marketing Res.* **8**(February) 20–29.

Rethans, A. J., J. L. Swasy, L. J. Marks. 1986. The effects of TV commercial repetition, receiver knowledge, and commercial length: A test of the two-factor model. *J. Marketing Res.* **23** 50–61.

Rossi, P., G. Allenby, R. McCullogh. 2005. Bayesian statistics and marketing. *Series in Probability and Statistics*. Wiley, Hoboken, NJ.

Silk, A. J., T. G. Vavra. 1974. The influence of advertising's affective qualities on consumer response. G. D. Hughes, M. O. Ray, eds. *Buyer/Consumer Information Processing*. University of North Carolina Press, Chapel Hill, NC.

Simon, H. 1982. ADPULS: An advertising model with wearout and pulsation. *J. Marketing Res.* **19**(August) 352–363.

Simon, J. L. 1965. A simple model for determining advertising appropriations. *J. Marketing Res.* **2**(August) 285–292.

Simon, J. L. 1969. The effect of advertising on liquor brand sales. *J. Marketing Res.* **6**(August) 301–313.

Simon, J. L. 1970. *Issues in the Economics of Advertising*. University of Illinois Press, Urbana, IL.

Srinivasan, V., H. A. Weir. 1988. A direct aggregation approach to inferring microparameters of the Koyck advertising-sales relationship from macro data. *J. Marketing Res.* **25**(May) 145–156.

Strong, E. C. 1972. The effects of repetition in advertising: A field experiment. Doctoral dissertation, Graduate School of Business, Stanford University, Stanford, CA.

Telser, L. G. 1964. Advertising and competition. *J. Political Econom.* **72** 537–562.

Unnava, H. R., R. Burnkrant. 1991. Effects of repeating varied ad executions on brand name memory. *J. Marketing Res.* **28**(November) 406–416.

Vakratsas, D., T. Ambler. 1999. How advertising works: What do we really know? *J. Marketing* **63**(Janaury) 26–43.

Vakratsas, D., F. Feinberg, F. M. Bass, G. Kalyanaram. 2005. The shape of advertising response functions revisited: A model of advertising effects with dynamic probabilistic thresholds. *Marketing Sci.* **23** 109–119.

Van Heerde, H. J., C. Mela, P. Manchanda. 2004. The dynamic effect of innovation on market structure. *J. Marketing Res.* **41**(2) 166–184.

Vidale, M. L., H. B. Wolfe. 1957. An operations research study of sales response to advertising. *Oper. Res.* **5**(June) 370–381.

Villas-Boas, M. J., R. Winer. 1999. Endogeneity in brand choice models. *Management Sci.* **45**(10) 1324–1338.

Weilbacher, W. M. 1970. What happens to advertisements when they grow up? *Public Opinion Quart.* **34**(Summer) 216–223.

West, M., J. Harrison. 1997. *Bayesian Forecasting and Dynamic Models*. Springer, New York.

Winer, R. S. 1979. An analysis of the time varying effects of advertising: The case of Lydia Pinkham. *J. Bus.* **52**(October) 563–576.

Winer, R. S., W. L. Moore. 1989. The effects of advertising and other marketing mix variables on brand positioning. *J. Advertising Res.* **28**(February/March) 39–45.

Wittink, D. R. 1977. Exploring territorial differences in the relationship between marketing variables. *J. Marketing Res.* **14**(May) 145–155.