

Planning Media Schedules in the Presence of Dynamic Advertising Quality

Prasad A. Naik • Murali K. Mantrala • Alan G. Sawyer

Graduate School of Management, A.O.B. IV, Room 151, University of California-Davis, Davis, California 95616

panaik@ucdavis.edu

Marketing Department, University of Florida, 207 Bryan Hall, Gainesville, Florida 32611

mantrala@dale.cba.ufl.edu

Marketing Department, University of Florida, 300A Bryan Hall, Gainesville, Florida 32611

sawyerag@dale.cba.ufl.edu

Abstract

A key task of advertising media planners is to determine the best media schedule of advertising exposures for a certain budget. Conceptually, the planner could choose to do continuous advertising (i.e., schedule ad exposures evenly over all weeks) or follow a strategy of pulsing (i.e., advertise in some weeks of the year and not at other times). Previous theoretical analyses have shown that continuous advertising is optimal for nearly all situations. However, pulsing schedules are very common in practice. Either the practice of pulsing is inappropriate or extant models have not adequately conceptualized the effects of advertising spending over time.

This paper offers a model that shows pulsing strategies can generate greater total awareness than the continuous advertising when the effectiveness of advertisement (i.e., ad quality) varies over time. Specifically, ad quality declines because of advertising wearout during periods of continuous advertising and it restores, due to forgetting effects, during periods of no advertising. Such dynamics make it worthwhile for advertisers to stop advertising when ad quality becomes very low and wait for ad quality to restore before starting the next "burst" again, as is common in practice.

Based on the extensive behavioral research on advertising repetition and advertising wearout, we extend the classical Nerlove and Arrow (1962) model by incorporating the notions of repetition wearout, copy wearout, and ad quality restoration. *Repetition wearout* is a result of excessive frequency because ad viewers perceive that there is nothing new to be gained from processing the ad, they withdraw their attention, or they become unmotivated to react to advertising information. *Copy wearout* refers to the decline in ad quality due to passage of time independent of the level of frequency. *Ad quality restoration* is the enhancement of ad quality during media hiatus as a consequence of viewers forgetting the details of the advertised messages, thus making ads appear "like new" when reintroduced later.

The proposed model has the property that, when wearout effects are present, a strategy of pulsing is superior to continuous advertising even when the advertising response function is concave. This is illustrated by a numerical example that compares the total awareness generated by a single concentrated pulse of varying duration (blitz schedules)

and continuous advertising (the even schedule). This property can be explained by the tension between the pressure to spend the fixed media budget quickly to avoid copy wearout and the opposing pressure to spread out the media spending over time to mitigate repetition wearout.

The proposed model is empirically tested by using brand-level data from two advertising awareness tracking studies that also include the actual spending schedules. The first data set is for a major cereal brand, while the other is for a brand of milk chocolate. Such advertising tracking studies are now a common and popular means for evaluating advertising effectiveness in many markets (e.g., Millward Brown, MarketMind).

In the empirical tests, the model parameters are estimated by using the Kalman filter procedure, which is eminently suited for dynamic models because it attends to the intertemporal dependencies in awareness build-up and decay via the use of conditional densities. The estimated parameters are statistically significant, have the expected signs, and are meaningful from both theoretical and managerial viewpoints. The proposed model fits both the data sets rather well and better than several well-known advertising models, namely, the Vidale-Wolfe, Brandaid, Litmus, and Tracker models, but not decisively better than the Nerlove-Arrow model. However, unlike the Nerlove-Arrow model, the proposed model yields different total awareness for different strategies of spending the same fixed budget, thus allowing media planners to discriminate among several media schedules.

Given the empirical support for the model, the paper presents an implementable approach for utilizing it to evaluate large numbers of alternative media schedules and determine the best set of media schedules for consideration in media planning. This approach is based on an algorithm that combines a genetic algorithm with the Kalman filter procedure. The paper presents the results of applying this approach in the case studies of the cereal and milk chocolate brands. The form of the best advertising spending strategies in each case was a pulsing strategy, and there were many schedules that were an improvement over the media schedule actually used in each campaign.

(Advertising Strategy; Advertising Wearout; Aggregate Response Models; Pulsing Schedules; Kalman Filter Estimation; Genetic Algorithm)

1. Introduction

Companies spend millions of dollars annually on advertising. Such advertising spending decisions for a specific brand are usually made in two stages. First, in the budgeting stage, the firm's management decides how much to spend on advertising and then, in the planning stage, the advertising agency of the firm recommends a media plan on how to spend the given budget during the upcoming year. These decisions include the questions of which media to use and when to schedule advertising over the year. This paper focuses on the scheduling decision. For example, the budget can be allocated evenly over the weeks of the year (an *even* spending strategy) or the firm may advertise for some weeks and not advertise at all in other weeks of the year (a *pulsing* strategy). If it is decided to consider a pulsing schedule, there are, of course, numerous variations of possible schedules. A media planner faces the problem of determining the schedule of ad exposures (that can be bought with the given budget) that will generate the largest total awareness for the advertised brand in the target market. This paper offers an implementable model-based approach for solving this media planning problem.

To address these issues, we propose a new aggregate advertising response model that incorporates the notion that effectiveness (or quality) of advertising copy varies over time, depending on whether advertising is exposed in a given time period. Our model offers a rationale for why a pulsing schedule may be more effective than the even one in some instances. Advertising quality may decline or wear out during periods of continuous advertising and it restores, due to forgetting effects, during periods of no advertising. It can be worthwhile to stop advertising when advertising quality gets very low and wait for it to restore before starting the next burst of advertising. In situations where wearout effects are negligible, these dynamics would be absent and then, consistent with the existing literature (e.g., Sasiemi 1989), a continuous spending strategy is optimal in the presence of concave (or linear) response functions. On the other hand, if wearout effects are significant, then advertisers benefit from using pulsing media schedules.

This paper is organized as follows: The next section reviews the empirical advertising research literature

that suggests that advertising effectiveness is subject to wearout and restoration effects followed by a review of extant advertising scheduling models. The proposed dynamic response model is developed in the third section, and empirical support for it is provided in the fourth section. This empirical analysis employs continuous advertising-awareness tracking data collected during two advertising campaigns for frequently purchased brands of packaged goods in England. Then, the fifth section develops an approach for determining the optimal media schedule incorporating the proposed response model and demonstrates its application in the cases of the two packaged-good brands. The paper concludes with a summary of its contributions and directions for future research.

2. Literature Review

In this section, we first review studies in the empirical advertising, consumer behavior, and marketing science literature that provide insights into how and why the effectiveness of advertising is time-varying. We then review previous efforts to develop advertising scheduling models and the directions for improving these models that we see in the light of the observations in the empirical literature.

2.1. Advertising Wearout

In advertising research, the term *wearout* refers to a decline in the quality or effectiveness of an advertisement (e.g., Grass and Wallace 1969; Greenberg and Suttoni 1973; Calder and Sternthal 1980; Pechmann and Stewart 1990). It is important to recognize the distinction between the effects of advertising wearout and those due to forgetting. Advertising wearout lowers the effectiveness of advertising copy (e.g., Pekelman and Sethi 1978), whereas forgetting leads to a decrease in aggregate brand awareness (e.g., Mahajan, Muller, and Sharma 1984). Thus, wearout refers to the effects of advertising when it is exposed, while forgetting refers to the degree of decline of the memory of advertising when the advertising is not being exposed. Broadly speaking, there are two sources of advertising wearout: *repetition wearout* and *copy wearout*.

Repetition wearout occurs as a result of excessive frequency of advertising and is a well-recognized problem in advertising practice (Corkindale and

Newall 1978). Although repeated advertising can produce many positive effects on such communication dimensions as recall, attitude, and purchase intention (e.g., Sawyer 1981), it is typical that repeated ad exposures eventually result in repetition wearout (see reviews by Greenberg and Suttoni 1973; Axelrod 1980; and Pechmann and Stewart 1990). Repetition wearout results when the potential ad viewers perceive that there is nothing new (such as added information or entertainment) to be gained from processing the advertisement (Weilbacher 1970; Hughes 1992) and, hence, become bored (Berlyne 1970; Cacioppo and Petty 1979) or irritated (Greysen 1973; Belch 1982). Because of these factors, consumers tend to withdraw their attention after several exposures (Grass and Wallace 1969; Krugman 1972), or, even with attention, become unmotivated to actively process and react to the advertising information (Cacioppo and Petty 1979; Calder and Sternthal 1980). This lack of active processing inhibits later recall of the ad, the advertised brand, and the information in the advertisement (Appel 1971; Craig, Sternthal, and Leavitt 1976), attitudes (Axelrod 1980; Blair 1988), and subsequent sales (Blair and Rosenberg 1994; Lodish et al. 1995).

Copy wearout refers to the decline in advertising quality due to the passage of time independent of repetitiveness or the level of frequency. Several causes for copy wearout have been identified. First, the conditions such as consumer knowledge or preferred attributes that justify a given copy or approach may change over time (e.g., Calantone and Sawyer 1978). For example, Assmus, Farley, and Lehmann (1984) note that when the ad copy is new it provides information on search attributes, but over time, as consumers acquire experience with the brand, the impact of the advertising copy dilutes. Also, a copy approach may wear out because "copycatting" takes place. That is, advertising style gets imitated by other firms, thereby lowering the perceived contrast among campaigns and intensifying the competition for consumers' attention (Axelrod 1980; Groenhaug, Kvitastein, and Gronmo 1991). For example, in just a single month, several brands in different product categories successively came up with the following claims: "All Fiber Is Not Created Equal" (Metamucil), "All Calories Are Not Created Equal" (Campbell's Soup), "All Gold Is

Not Created Equal" (Visa), and "All Cigarettes Are Not Created Equal" (Kool). Finally, there may be changes over time in the amount of "clutter" from competitive and other advertisements, which increases the "noise" level in the environment and raises the threshold of the minimum number of exposures required to impact the consumer (Corkindale and Newall 1978; Webb and Ray 1979; Groenhaug, Kvitastein, and Gronmo 1991).

2.2. Advertising Quality Restoration

Several researchers have noted that advertising wearout can be forestalled and advertising quality restored by using an appropriate media scheduling strategy. Practitioners believe that spacing between bursts of advertising is beneficial. For instance, Grass and Wallace (1969, p. 8) note that "regeneration of attention or interest level (in ads) is possible . . . if they are removed from the air." Similarly, Greenberg and Suttoni (1973, p. 53) point out that "a commercial that is running for a while can be removed and reintroduced after a time and [it can] take on a sense of newness." Some experimental and field studies have supported these observations. Calder and Sternthal's (1980) experimental results show that spacing between bursts of advertising influences the perceived arousal of commercials as well the amount of cognitive responses. A field study by Grass and Wallace (1969) indicates that a gap or hiatus between media bursts enhances the attention to advertisements. A reason for restoration in ad quality during a media hiatus is that consumers forget the details of advertised messages and these ads appear "like new" when reintroduced later on; hence, the greater the forgetting, the more restored is the ad quality when advertising is resumed (Corkindale and Newall 1978, p. 334).

2.3. Media Scheduling Models

Strong (1977), Zielske and Henry (1980), and Katz (1980) attempted to develop model-based decision aids for media scheduling. These models used simulations based on the empirical results of two field experiments (Strong 1974; Zielske 1959).¹ These models were not particularly successful and provided no way to suggest an optimal schedule or to definitively decide

¹See also the comment of Simon (1979) on Zielske's (1959) results.

whether a continuous or a pulsing schedule was more effective. Subsequently, it was elegantly proved by Hartl (1987) that a single-state variable (i.e., lagged awareness) model such as these could never yield a pulsing strategy as optimal.

Other research on media scheduling has been largely theoretical in nature with the focus on identifying conditions under which different types of dynamic media scheduling strategies are optimal. Curiously, except for Pekelman and Sethi (1978), none of these models has discussed or included advertising wearout (see, e.g., Feichtinger, Hartl, and Sethi 1994).² It has been shown that a continuous spending strategy is optimal when brand awareness or sales is a concave function of advertising spending, while a very rapid switching between “on” and “off” advertising (termed a *chattering* strategy) is optimal when the function has an S shape (e.g., Sasieni 1971; Mahajan and Muller 1986). However, a chattering strategy is not implementable in practice and is perceptually indistinguishable from the continuous one. This was pointed out by Feinberg (1992), who also showed that a pulsing media schedule with a finite duration of advertising bursts is superior to chattering as well as continuous strategies when an S-shaped response function is combined with an exponential filtering mechanism. Given the limited empirical support for S-shaped response functions (e.g., Rao and Miller 1975; Simon and Arndt 1980), this body of research suggests that continuous advertising schedules should be prevalent in practice.

Simon (1982) offers a model that allows a pulsing strategy to dominate the continuous strategy when sales response to advertising is not S-shaped. In this model, the superiority of pulsing is driven by the asymmetry of sales response. Specifically, the gain in sales volume when advertising spending is increased by a certain amount is greater than the loss in sales

when the spending is reduced by the same amount. In such situations, alternating between high and low spending levels makes sense. However, as noted by Sasieni (1989), there is only limited empirical support for this model specification (Haley 1978). Further, both Simon (1982) and Mesak (1992) note that their models leave unresolved important issues concerning the duration of and the spacing between advertising bursts.

2.4. Implications for Model Building

The phenomena of advertising wearout and restoration during a hiatus imply that the advertising effectiveness parameter in aggregate advertising response models should be allowed to vary over time. Consistent with this view, Parsons and Schultz (1976, p. 85) advocate the use of a “quality adjustment factor” over and beyond the amount or quantity of media spending. In other words, the quality adjustment factor is a time-varying coefficient for the effectiveness of advertising expenditure in models of awareness formation. Arnold et al. (1987) provide empirical support for this formulation of advertising quality. Given that advertising wears out, it is reasonable that a model should incorporate such effects into its formulation. Also, given the Hartl (1987) proof, it seems appropriate to have a model with at least two states (advertising quality as well as awareness) to allow for the possibility that a pulsing schedule would be superior to an even one, rather than a single-state model with no opportunity to test whether a pulsing schedule might be optimal. In the next two sections, we build on this work by specifying the dynamics of ad quality due to wearout and restoration and empirically testing this specification. In § 5, we examine the implications of these dynamics for planning pulsing media schedules.

3. Model Development

3.1. Model of Awareness Formation with Dynamic Advertising Quality

The marketing literature contains a variety of formulations of aggregate advertising response models (see, e.g., Little 1979). The most parsimonious of these is the classic goodwill accumulation model of Nerlove and Arrow (1962). We choose to use the Nerlove-Arrow

²Simon (1982) and Mesak (1992) use the term advertising wearout to describe the phenomenon of asymmetric response to increases and decreases in advertising spending. A better term for this phenomenon would perhaps be *hysteresis*, as suggested by Sasieni (1989), since the path traced by a dependent variable (sales) when advertising is increased is not the same as its path when advertising is decreased. Our usage of the term advertising wearout is consistent with the advertising and consumer behavior literatures (e.g., Greenberg and Suttoni 1973; Calder and Sternthal 1980).

(N-A) model as a basis for our proposed model development not only because of its parsimony but also because this representation of aggregate advertising response has been shown to be consistent with a theoretically well-grounded “micro-model” of individual consumer response to advertising (Blattberg and Jeuland 1981; Rao 1986). As pointed out by Little (1979) and Mahajan and Muller (1986), the N-A model can be used to relate advertising to awareness by considering awareness as a form of goodwill.³

More specifically, the N-A model is:

$$\frac{dA}{dt} = qu(t) - \delta A, \quad (1)$$

where $A(t)$ is the awareness at time t , δ is the decay or forgetting rate, $u(t)$ is the advertising spending rate (e.g., dollars or gross rating points (GRPs) per week), and q is a constant advertising quality or effectiveness coefficient, assumed to be equal to one by Nerlove and Arrow (1962). Equation (1) implies the awareness growth rate dA/dt increases linearly with the media spending rate $u(t)$, and decreases due to forgetting, which is proportional to the level of awareness. However, based on our review of the empirical advertising research literature, we propose to extend the N-A model to take into account the dynamics of advertising quality.

Specifically, we posit that instead of being constant, advertising quality $q = q(t, u(t))$. That is, ad quality is a function of t , the time that has elapsed since the onset of the campaign, and $u(t)$, the time pattern of advertising spending. More precisely, we assume, like Pekelman and Sethi (1978), that advertising quality or effectiveness decays with the passage of time (the copy wearout effect). We also assume that advertising quality is negatively impacted by the advertising intensity (spending level) when advertising is underway (the repetition wearout effect) but restores due to forgetting when advertising is stopped. Thus, our proposed

advertising-awareness response model is defined by two differential equations—one representing the awareness evolution, and the second representing advertising quality evolution over time. These equations of the proposed model are given below followed by an explanation of each equation’s parameters and its underlying assumptions:

$$\frac{dA}{dt} = qg(u) - \delta A \quad (2)$$

and

$$\frac{dq}{dt} = -a(u)q + (1 - I(u))\delta(1 - q), \quad (3)$$

where $I(u)$ is such that,

$$I(u) = \begin{cases} 1 & \text{if } u \neq 0, \\ 0 & \text{if } u = 0, \end{cases} \quad (4)$$

and

$$a(u) = c + wu(t), \quad (5)$$

where c and w denote the copy wearout and repetition wearout parameters, respectively.

Equation (2) is the awareness formation model in which $g(u)$ denotes the instantaneous advertising response function such that $g(u) \geq 0$ for $u \geq 0$. The form of $g(u)$ may be specified to be linear ($g(u) = u$) as in the Nerlove and Arrow (1962) model, concave as in Gould (1970), or possibly S-shaped as in Mahajan and Muller (1986). The term $qg(u(t))$ denotes the instantaneous “advertising impact” (e.g., Parsons and Schultz 1976; Arnold et al. 1987). In other words, the main effect of advertising spending $u(t)$ on the awareness growth rate is moderated by the ad quality q at time t . Like the original N-A model, it is assumed that the maximum value of q is scaled to be equal to one. Thus, if $g(u)$ is assumed to be linear and q is assumed to be a constant, Equation (2) reduces to the Nerlove-Arrow model.

Equation (3) represents the dynamics of advertising quality assumed by the proposed model. Specifically, when $I(u) = 1$, i.e., advertising is “on,” advertising quality decays due to the two types of wearout according to the differential equation

$$\frac{dq}{dt} = -a(u)q. \quad (6)$$

³We acknowledge that awareness of a brand in a target market can be influenced by several marketing variables in addition to advertising, e.g., free samples, in-store displays, etc. However, in the following model development, we restrict our scope only to advertising as we are interested specifically in developing a model useful for a company’s advertising media schedule planning and decision making.

That is, we assume, first, that when $c > 0$ and/or $w > 0$, advertising quality exponentially decays when advertising is on.⁴ Second, consistent with the literature review in § 2.1, we assume that advertising quality decay due to copy wearout continues even when advertising spending has been suspended (i.e., $a(0) = c$). Third, the advertising quality decay rate is proportional to the level of spending (or repetition⁵) where the repetition wearout parameter w captures the marginal contribution of spending to the overall advertising wearout (i.e., $\partial a / \partial u = w$). Fourth, if both wearout effects are absent, i.e., $c = 0$ and $w = 0$, then advertising quality remains constant when advertising is on, as in the N-A model.

Finally, when advertising is turned “off” after it has been initiated, i.e., $I(u) = 0$, Equation (3) reduces to the differential equation

$$\frac{dq}{dt} = -cq + \delta(1 - q). \quad (7)$$

Thus, we assume that while advertising quality continues to decline due to copy wearout, it can also restore during a hiatus in advertising spending as a result of forgetting. This follows from the literature review in § 2.2. Of course, repetition wearout wanes when advertising is stopped. Further, consistent with the assertions of Corkindale and Newall (1978, p. 334), we assume that the regeneration rate of advertising quality is proportional to the forgetting rate δ . The greater the forgetting rate, the more the enhancement of ad quality. Lastly, after advertising has stopped, ad quality restores, and it reaches the limiting value

$$\lim_{\tau \rightarrow \infty} q(\tau) = \frac{\delta}{c + \delta} \leq 1,$$

which is equal to unity only when copy wearout is negligible, i.e.,

⁴Pekelman and Sethi (1978) have used a linear decay of ad quality (i.e., $dq/dt = -a$) to model ad wearout, while we use a linear decay of log(ad quality), i.e., $d \ln q / dt = -a(u)$. The benefit of the log transformation is that we ensure that $q(t) > 0$ for all t , as it should be.

⁵A higher spending rate implies higher repetition because, for simplicity, we assume that a single advertisement copy is run in a given medium that has a fixed reach and cost per thousand during the planning horizon. Also, note the discussion of the appropriateness of this assumption in data description in § 4.1.

$$\lim_{\tau \rightarrow \infty} q(\tau) = 1 \quad \text{as } c \rightarrow 0.$$

That is, advertising quality can never be restored to being “as good as new” if copy wearout is significant.

3.2. An Illustration of the Model’s Theoretical Properties: Should Advertising Be Concentrated?

In § 2.4, we identified the need to develop a plausible aggregate advertising response model that can be used by media planners to test whether a pulsing schedule might be optimal, whatever the shape of the response function may be. We now provide an illustration of how our proposed model meets this need.⁶ In this illustration, we utilize the proposed model to evaluate and compare the performance of a special case of advertising pulsing, namely, spending the total budget in a single pulse or *blitz* (see Mahajan and Muller 1986), with that of an even spending policy. This illustration assumes a concave specification of the instantaneous advertising response function and specific numerical values of the model parameters.

More specifically, consider a blitz schedule with expenditure of the total available advertising media budget B concentrated in the first l weeks of a planning horizon of length $T > l$ weeks as follows:

$$\begin{aligned} u(t) &= \bar{u}, & t \in [0, l], \\ &= 0, & t \in [l, T], \end{aligned} \quad (8)$$

i.e., a constant spending rate $\bar{u} = B/l$ is maintained when advertising is on.

Now, let the ad response function $g(\bar{u})$ be concave with the specification $g(\bar{u}) = \log(1 + \bar{u})$, as in Luhmer et al. (1988). Further, let $c = 0.1$, $w = 0.01$, $\delta = 0.001$, $B = 100$, $T = 52$ weeks, and the initial awareness level $A_0 = 0$. For these parameter values and $\bar{u} = B/l$, the solutions to the proposed model Equations (2)–(5) can

⁶Simon’s (1982) model does yield pulsing as an optimal policy without an S-shaped response function. However, this model has very little empirical support and is theoretically unappealing because a pulsing schedule will always dominate an even schedule under this model (see, e.g., Mahajan and Muller 1986). In contrast, our proposed model allows for wearout effects but does not necessarily imply that a pulsing schedule is inevitably optimal. The optimality of either type of schedule depends on the estimated parameters of our model in any application.

be used to compute total awareness J for various values of l . The expressions involved in these computations are provided in Appendix A. Note that when $l = T$, we effectively have a blitz lasting the whole year, which is nothing but the even schedule. Figure 1 depicts the graph of $J(l)$ over l weeks.

Figure 1 shows that the optimal blitz schedule expends the entire media budget in 19 weeks, and this single pulse schedule generates the maximum total awareness of 11,359 units. Since the even schedule, which expends the media budget in 52 weeks, generates only 9,000 units of total awareness, it is not the optimal schedule. Moreover, several blitz schedules, those which concentrate budget spending over 10 weeks to 51 weeks, perform better than the even schedule. However, we note that the model does not imply that blitz schedules necessarily outperform the even schedule in the presence of ad wearout. As can be seen in Figure 1, the even schedule is superior to many highly concentrated blitz schedules, such as those that concentrate all the spending in six or fewer weeks.

The intuition for these results is the tension between the effects of copy and repetition wearout. Copy wearout exerts pressure to spend as much as possible of the media budget right at the beginning of the planning horizon when ad quality is still high, rather than spread the budget over the entire horizon. However, this implies a high intensity of spending (high repetition of exposures), which exacerbates repetition wearout and thus exerts pressure to spread out the media spending, rather than concentrate, over the horizon. These two opposing forces lead to an interior solution

for the optimal duration of blitz, l^* , as illustrated in the numerical example.

To summarize, this illustration shows that even when the advertising response function is concave, a pulsing schedule can be superior to the even spending schedule due to the opposing pressures of copy and repetition wearout effects incorporated in our model. In the absence of these wearout effects, media scheduling is simple: Spend the budget uniformly throughout the planning horizon. However, our model shows a pulsing strategy is optimal if wearout effects are significant. This property of the proposed model is consistent with observations in the literature that advertisements do wear out (e.g., Blair 1988) and media schedules are often pulsed in practice (e.g., Corkindale and Newall 1978; Jones 1995). The latter observation has led Little (1986) to ask, "Are there any response models for which pulsing (other than chattering) would be optimal?" Our illustration suggests that the proposed model is a plausible formulation that does indeed have this property and can be utilized for planning media schedules. To substantiate these points further, we now proceed to estimate and validate the proposed model utilizing data from the actual campaigns for two brands that have been reported in the previous literature.

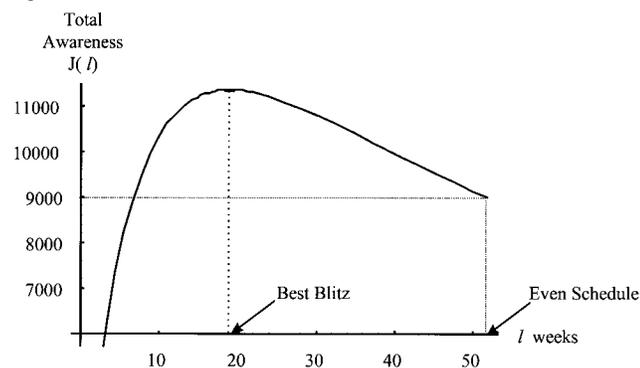
4. Model Estimation

In this section we use advertising-awareness data collected during the tracking studies of advertising campaigns for a cereal brand and a milk chocolate brand in the UK. We begin with the cereal brand case study. We first describe the general data collection and the estimation procedures, and then present estimation results, robustness to alternative model specifications, and, finally, model comparisons with several extant aggregate response models. This is followed by a description and a summary of model estimation results obtained with the milk chocolate brand data.

4.1. Cereal Brand Advertising Case Study

4.1.1. DESCRIPTION OF THE DATA. The cereal brand advertising data set comprises a weekly time series of the advertising levels (measured in England by TVRs, which are highly similar to gross rating

Figure 1 Total Awareness as a Function of Duration of Blitz



points in the United States) and advertising-awareness scores; these data were taken from research published by West and Harrison (1997).

Figure 2 depicts the actual media spending schedule employed by the cereal company, and Figure 3 shows the plot of weekly ad awareness. The same single campaign was telecast during the 75-week campaign period. The pattern of spending in Figure 2 indicates that the cereal company employed a pulsing media schedule. In particular, we note the long media hiatus during Weeks 9 through 16 and Weeks 49 through 75, as well as other shorter nonspending gaps in the intervening weeks.

These data are appropriate to test our model for the following reasons. First, advertising tracking data are widely used by large advertisers and are considered to be useful in determining advertising effectiveness (Rossiter and Percy 1997). Advertising awareness has been shown to relate positively to sales in some cases (see, for example, Colman and Brown 1983 for an analysis of the relationship of these data to sales). Second,

advertising awareness is likely to reflect the prior noting and processing of the advertising, such as we conceptualized earlier. Third, the advertising campaign includes only a single ad copy, which is an assumption of our model. Fourth, the fact that the data are for television advertising in England in the late 1970s is a strength, since the TVR estimates of advertising exposure are a reasonable proxy for the frequency of exposure to TV advertising. This is because an assumption of constant reach for each exposure and a given amount of spending holds reasonably well with these data because there were only two television channels in England showing advertisements at that time (Brown 1985).

The data were collected in in-home surveys administered by Millward Brown. Weekly personal interviews were administered to 66 households on average. The measure of advertising effectiveness was advertising recall or awareness, which is measured with the following question after showing a list of brands: "Which of these brands of _____ have you seen advertised on television recently?" Advertising awareness is measured separately for each medium.

Figure 2 Actual TV Advertising Schedule for the Cereal Brand

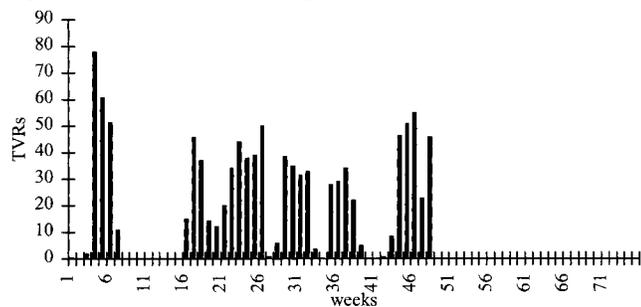
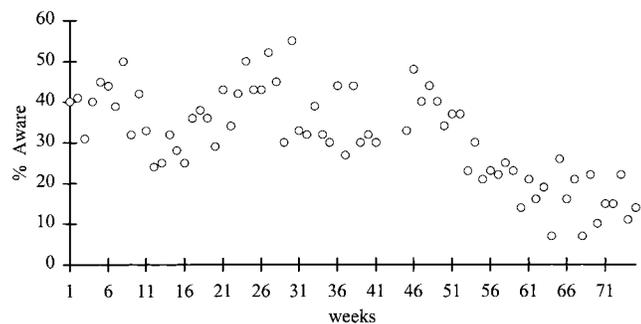


Figure 3 Percentage of Respondents Aware of Cereal Brand's TV Advertising



4.1.2. KALMAN FILTER ESTIMATION. The model parameters are estimated by using the Kalman filter methodology (see Harvey 1994). This approach is eminently suited for the dynamic model specified by Equations (2)–(5) because it takes into account the dynamic evolution of both the state variables, awareness and ad quality, simultaneously. This approach consists of three stages. In the first stage, we obtain a *transition equation* from the dynamic model for these state variables. The second stage links the transition equation to the tracking data via an *observation equation*. Finally, to obtain the parameter estimates, the *conditional likelihood function* is derived by using the transition and observation equations. The interested reader can obtain further details on how the Kalman filter works in Harvey (1994) or Hamilton (1994, Ch. 13). Below, we describe the three stages.

Obtaining the Transition Equation. Since the tracking data are collected at discrete points of time (weeks), we discretize Equations (2) and (3) and re-express them in the following vector form:

$$\begin{aligned} \begin{bmatrix} A_t \\ q_t \end{bmatrix} &= \begin{bmatrix} (1 - \delta) & & & \\ 0 & (1 - a(u)) - \delta(1 - I(u)) & & \\ & & g(u_t) & \\ & & & \delta(1 - I(u)) \end{bmatrix} \begin{bmatrix} A_{t-1} \\ q_{t-1} \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ \delta(1 - I(u)) \end{bmatrix} + \begin{bmatrix} \nu_{1t} \\ \nu_{2t} \end{bmatrix}. \end{aligned} \quad (9)$$

We have introduced the error terms $\nu_t = [\nu_{1t}, \nu_{2t}]'$ to capture the net effects of myriad variables that potentially affect the dynamics of awareness and ad quality that cannot be modeled explicitly for the sake of parsimony. Equation (9) is the transition equation, which can be compactly expressed in vector notation as $\alpha_t = T_t \alpha_{t-1} + c_t + \nu_t$, where $\alpha_t = [A_t, q_t]'$ is a 2×1 state vector, T_t is a 2×2 transition matrix, and c_t is a 2×1 drift vector.

Linking the Transition Equation to Observed Awareness Scores. Let Y_t denote the observed awareness score obtained from the tracking surveys in week t . Clearly, the observed awareness scores Y_t will not exactly equal the model-based awareness A_t . So we have the relation $Y_t = A_t + \epsilon_t$, where ϵ_t denotes the measurement errors due to the sampling process of surveys. Since the dimensions of the observed Y_t and the state vector are different, we re-express the relation as

$$Y_t = [1 \ 0] \begin{bmatrix} A_t \\ q_t \end{bmatrix} + \epsilon_t. \quad (10)$$

Equation (10) is termed the observation equation, which can be compactly expressed as $Y_t = z \alpha_t + \epsilon_t$, where $z = [1 \ 0]$ is a vector of constants. The transition and observation equations together constitute the *state-space form*.

Deriving the Conditional Likelihood Function. To derive the likelihood function, let $p(Y; \Theta)$ denote the probability of jointly observing the sequence of awareness scores $Y = (Y_1, \dots, Y_T)'$, where Θ is a vector of unknown parameters. We decompose this joint density as the product of conditional density times the marginal density. That is,

$$p(Y; \Theta) = f(Y_T | \mathcal{T}_{T-1}) \times p(Y_1, \dots, Y_{T-1}; \Theta),$$

where $f(Y_T | \mathcal{T}_{T-1})$ denotes the conditional density of observing the awareness score, Y_T , given the information set \mathcal{T}_{T-1} . This set contains all information on awareness and media spending thus far, except the current period awareness score; that is, $\mathcal{T}_t = \{Y_1, \dots, Y_{t-1}; u_1, \dots, u_t\}$. Applying the decomposition of the

joint density repeatedly, we write the conditional likelihood function, $L(\Theta; Y)$, as follows:

$$\begin{aligned} L(\Theta; Y) &= p(Y; \Theta) \\ &= f(Y_T | \{Y_1, \dots, Y_{T-1}\}) p(Y_1, \dots, Y_{T-1}; \Theta) \\ &= f(Y_T | \mathcal{T}_{T-1}) f(Y_{T-1} | \{Y_1, \dots, Y_{T-1}\}) \\ &\quad p(Y_1, \dots, Y_{T-2}; \Theta) \\ &= \dots \\ &= \prod_{t=1}^T f(Y_t | \mathcal{T}_{t-1}), \quad t = 1, \dots, T. \end{aligned} \quad (11)$$

Parameter Estimation and Statistical Significance. To obtain parameter estimates, we need the moments of the conditional density $f(Y_t | \mathcal{T}_{t-1})$. The moments are obtained by assuming that error terms $(\epsilon_t, \nu_{1t}, \nu_{2t})'$ are independently and identically distributed as normal random variables with zero means. Specifically, we have,

$$[\epsilon_t \ \nu_{1t} \ \nu_{2t}]' \sim \text{MVN} \left(0_{3 \times 1}, \begin{bmatrix} H_{1 \times 1} & 0 \\ 0 & Q_{2 \times 2} \end{bmatrix} \right). \quad (12)$$

Under the above assumption, Appendix B shows that $f(Y_t | \mathcal{T}_{t-1})$ is a normal density and provides the mean and variance of the random variable $Y_t | \mathcal{T}_{t-1}$ in closed form recursively for all t . It can be seen from Equation (B3) in Appendix B that the mean and variance are functions of observed awareness scores Y_t . Hence, the likelihood of observing a specific sequence of awareness scores can be evaluated by using Equation (11) for some trial parameter values. The estimated parameters, Θ^* , are those that maximize this likelihood value. To determine the significance of estimated parameters, the standard errors are obtained by evaluating the information matrix (Harvey 1994, p. 140) at the estimated parameter values. Since the parameter estimates are normally distributed, the statistical significance is based on the asymptotic t test, as in the standard regression analysis.

The above Kalman filter uses conditional density, rather than marginal, in constructing the likelihood function to obtain parameter estimates. Consequently, the intertemporal dependence in awareness scores induced by the model dynamics (i.e., Equations (2)–(5)) is retained. This will not be the case if marginal density

were used, as in the ordinary or nonlinear least squares approaches. Therefore, the parameter estimates obtained from the Kalman filter methodology have desirable statistical properties. Specifically, they are consistent, asymptotically unbiased, and have the minimum variance among a class of any estimators (Harvey 1994).

4.1.3. CEREAL DATA-BASED MODEL ESTIMATION RESULTS. Using the above approach, the proposed model was estimated assuming three different specifications of $g(u)$, namely, $g(u) = u$, $g(u) = \sqrt{u}$ and $g(u) = \ln(1 + u)$. The maximized log-likelihood values for the three specifications were -172.32 , -174.41 , and -175.64 , respectively. Hence, we retain the simpler linear specification rather than square-root or log transformations. To locate the globally optimal parameter estimates, several estimation runs were performed with widely different starting parameter values. The estimation runs converged to the same parameter estimates, which enhances our confidence in them. The estimate for the repetition wearout parameter, \hat{w} , was found to be negligible and not significant in all the estimation runs. Hence, the repetition wearout parameter was dropped from the final estimation run so that the precision (i.e., standard errors) of other parameters is correctly estimated (see Agresti 1990, p. 183). To keep the Kalman gain factor non-zero for updating the moments (see Equation (B3)), the transition noise parameters were retained in the estimation runs despite their small magnitude. Table 1 provides the parameter

estimates, standard errors, and t values obtained in the final estimation run of the proposed model with a linear ad response function.

In Table 1, the estimated parameters have the expected signs, are statistically significant, and are meaningful from statistical and managerial viewpoints. For example, the estimated initial awareness $\hat{A}_0 \approx 42\%$ is fairly close to the first few awareness scores recorded in the raw data (see Figure 3). Further, we see that measurement noise, $\hat{\sigma}_\epsilon = 5.82$, is a large and significant source of uncertainty—which is not surprising, given the small sample sizes in the tracking surveys. The forgetting rate parameter is also significant, consistent with what has been observed in previous studies (see, e.g., Mahajan, Muller, and Sharma 1984). Finally, the estimate of the copy wearout parameter ($\hat{c} = 0.123$) is relatively large and statistically significant ($t = 21.6$).

The significant estimates for the copy wearout and forgetting rate parameters suggest that ad quality for the cereal brand's advertising did vary over time. Figure 4 illustrates the time-varying effectiveness of the cereal brand's advertising. These results were obtained by substituting the estimated parameters \hat{c} and $\hat{\delta}$ in Equations (6) and (7). Thus, we note that there is a decline in ad quality when advertising is "on" due to copy wearout and a rise in ad quality when advertising is "off" due to the forgetting effect dominating the ongoing copy wearout effect. Given these dynamics of ad quality, it appears sensible for the cereal company to use a pulsing spending pattern in which advertising is discontinued when advertising quality gets quite low and resumed after a hiatus in spending during which the ad quality is restored.⁷

In addition to lending support to the hypothesized ad quality dynamics, the overall aggregate response model produces a good fit to the observed awareness data. This is illustrated in Figure 5.

4.1.4. MODEL COMPARISONS. In Table 2, we briefly describe the extant models we have chosen to

Table 1 Parameter Estimates for the Proposed Model Using Cereal Brand Data

Parameters	Estimates	Standard Error	T Value
Copy wearout, c	0.123	0.005	21.58
Forgetting rate, δ	0.043	0.004	10.03
Initial awareness, A_0	41.99	4.651	9.03
Initial ad quality, q_0	0.015	0.321	0.048
Measurement noise, σ_ϵ	5.819	0.466	12.49
Transition noise (awareness), $\sigma_{\epsilon 1}$	7.3×10^{-7}	0.515	1.4×10^{-6}
Transition noise (ad quality), $\sigma_{\epsilon 2}$	1.1×10^{-9}	0.003	3.3×10^{-7}
Maximized log-likelihood, LL*		-172.316	
Number of parameters		7	

⁷The above estimation results assume that the model residuals are normally distributed, are homoscedastic, have no autocorrelation, and the estimated parameters are stable (i.e., constant) over the time frame of the investigation. The validity of these assumptions was tested by performing model diagnostics in the time domain (see Harvey 1994, p. 256). Based on these tests, we verified that all the error term assumptions were upheld in this analysis.

Figure 4 Time-Varying Effectiveness of the Cereal Brand's Advertising

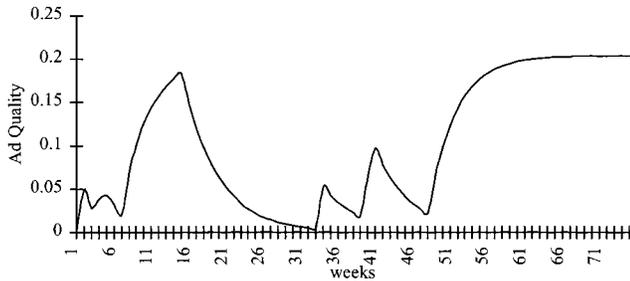
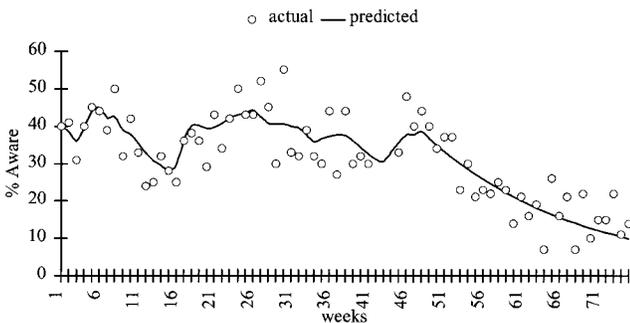


Figure 5 Model-Based Awareness Forecast Versus Observed Awareness Scores



compare with the proposed model, namely, the models of Vidale and Wolfe (1957), Nerlove and Arrow (1962), Brandaid (Little 1975), Tracker (Blattberg and Golanty 1978), and Litmus (Blackburn and Clancy 1982). It is interesting to note that these five well-known models can be cast into the same state-space form that was employed to estimate the proposed model. We recall that the state-space form comprises the observation equation, $Y_t = z \alpha_t + \epsilon_t$, and the transition equation, $\alpha_t = T_t \alpha_{t-1} + c_t + \nu_t$, with disturbance terms $\epsilon_t \sim N(0, H)$ and $\nu_t \sim N(0, Q)$. Table 3 shows the system matrices (e.g., T, H, Q) for the comparison models.

We estimated the five comparison models using the Kalman filter methodology. Table 4 presents the parameter estimates and asymptotic t values in parentheses. The estimated parameter values are meaningful as well as comparable across models. For example, all five models indicate initial awareness of about 40%, which is indeed close to the first observation in the data set. Similarly, the estimates of transition and measurement noises are similar across all models. As noted earlier (see Table 1), the measurement noise is the largest source of uncertainty.

Comparison of Model Fits. Table 4 gives the maximized log-likelihood values and the number of parameters of each model. To compare models, we use AIC

Table 2 Description of Comparison Models

Model	The Mathematical Model	Model Description
Vidale-Wolfe	$\frac{dA}{dt} = \beta(1 - A)u - \delta A$ <p>Discrete version</p> $A_t = (1 - \beta u_t - \delta)A_{t-1} + \beta u_t$	Over a small period of time, increase in brand awareness is due to the brand's advertising effort, which influences the unaware segment of the market, while attrition of the aware segment occurs due to forgetting of the advertised brand.
Nerlove-Arrow	$\frac{dA}{dt} = \beta u - \delta A$ <p>Discrete version</p> $A_t = (1 - \delta)A_{t-1} + \beta u_t$	The growth in awareness depends linearly on the advertising effort, while awareness decays due to forgetting of the advertised brand.
Brandaid	$A_t = \lambda A_{t-1} + (1 - \lambda)g(u_t)$ $g(u) = \frac{u^\phi}{\phi + u^\phi}$	Brand awareness in the current period depends partly on the last period brand awareness and partly on the response to advertising effort; the response to advertising effort can be linear, concave, or S-shaped.
Tracker	$A_t - A_{t-1} = (1 - e^{-\beta u_t})(1 - A_{t-1})$	The incremental awareness depends on the advertising effort, which influences the unaware segment of the market.
Litmus	$A_t = (1 - e^{-\beta u_t})A^* + e^{-\beta u_t}A_{t-1}$	The current period awareness is a weighted average of the steady-state ("maximum") awareness and the last period awareness. The weights are determined by the advertising effort in period t .

NAIK, MANTRALA, AND SAWYER
Planning Media Schedules

Table 3 System Matrices for Comparison Models

System Matrices	Vidale-Wolfe	Nerlove-Arrow	Brandaid	Tracker	Litmus
State vector, α_t	[A]	[A]	[A]	[A]	[A]
Observation vector, z	[1]	[1]	[1]	[1]	[1]
Transition matrix, T_t	$[1 - g(u) - \delta]$	$[1 - \delta]$	$[\lambda]$	$[1 - g(u)]$	$[1 - g(u)]$
Drift vector, c_t	$[g(u)]$	$[g(u)]$	$[(1 - \lambda)g(u)]$	$[g(u)]$	$[A^*g(u)]$
Observation noise, H	σ_ϵ^2	σ_ϵ^2	σ_ϵ^2	σ_ϵ^2	σ_ϵ^2
Transition noise, Q	σ_v^2	σ_v^2	σ_v^2	σ_v^2	σ_v^2
Response function, $g(x)$	βx	βx	$x/(\phi + x)$	$1 - e^{\alpha - \beta x}$	$1 - e^{-\beta x}$

Table 4 Parameter Estimates for Models Under Consideration

Parameters	Vidale-Wolfe	Nerlove-Arrow	Brandaid	Tracker	Litmus	Our Model
Response function $g(x)$:						
β	0.0 (0.0)	0.100 (5.2)	—	0.0 (0)	0.0 (0)	—
ϕ	—	—	0.953 (4.3)	—	—	—
γ	—	—	1.088 (2.6)	—	—	—
α	—	—	—	-0.012 (-1.3)	—	—
Copy wearout, c	—	—	—	—	—	0.123 (21.6)
Forgetting, δ	0.012 (1.4)	0.059 (5.4)	—	—	—	0.043 (10.0)
Smoothing, λ	—	—	0.988 (11.6)	—	—	—
Maximum awareness, A^*	—	—	—	—	98.4 (31.0)	—
Initial awareness, A_0	40.66 (7.7)	41.33 (9.0)	40.71 (9.3)	40.66 (7.7)	39.29 (7.5)	41.99 (9.03)
Initial ad quality, q_0	—	—	—	—	—	0.015 (0.048)
Measurement noise, σ_ϵ	5.851 (10.2)	5.656 (10.6)	5.851 (10.2)	5.850 (10.6)	5.820 (10.2)	5.819 (12.5)
Transition noise (awareness), σ_{v_1}	2.368 (4.1)	1.091 (2.0)	2.339 (4.1)	2.370 (4.2)	2.497 (4.5)	7.3×10^{-7} (1.4×10^{-6})
Transition noise (ad quality), σ_{v_2}	—	—	—	—	—	1.1×10^{-9} (3.3×10^{-7})
Maximized log-likelihood, LL*	-184.89	-172.99	-184.65	-184.89	-185.77	-172.316
Numbers of parameters	5	5	6	5	5	7
AIC = $-2LL^* + 2p$	379.78	355.98	381.3	379.78	381.54	358.64
BIC = $-2LL^* + p\text{Log}(75)$	391.37	367.57	395.2	391.37	393.13	374.86

(Akaike Information Criterion) and BIC (Bayes Information Criterion) metrics (see Harvey 1994, p. 80). Specifically, the former measure is defined as $AIC = -2(LL^*) + 2p$, where LL^* is the maximized log-likelihood value and p is the number of model parameters. Similarly, $BIC = -2(LL^*) + p\text{Log}T$, where T is the sample size. The terms $2p$ and $p\text{Log}T$ are interpreted as the *penalty* for adding more parameters. A smaller value for AIC and BIC is more desirable, i.e., it indicates a better model fit to the data. The last two

rows in Table 4 display the AIC and BIC measures for all six models under consideration.

According to the AIC and BIC measures, the goodness of fit of the proposed and N-A models is distinctly superior to that of the other four comparison models. Thus, we can reasonably drop these four models from further consideration. The AIC and BIC measures indicate that the N-A model fits these data marginally better than the proposed model. Specifically, the N-A model is better on the AIC metric by 0.75% and on the

BIC scale by 1.98%. Since so far it appears that there is little to choose between the proposed and N-A models, we probe this issue further by comparing the predictive performances of these two models.

4.1.5. COMPARING THE PROPOSED AND N-A MODELS.⁸ *Cross-Validation Analyses.* To compare predictive performances, we perform a cross-validation analysis that uses a part of the sample data to estimate model parameters and forecasts the remaining data by using the estimated model. The model forecasts are then compared with the actual outcomes in the holdout sample to compute the R^2 value, which is a measure of the predictive performance of the model. In cross-validation, the estimation sample sizes were 30, 31, 32, 33, and 34 weeks, and the holdout sample sizes were 45, 44, 43, 42, and 41 weeks, respectively. Figure 6 shows the predictive performance of the N-A and proposed models as a function of the estimation sample size. As a benchmark, we also depict the predictive performance of a simple random walk model that assumes the next period awareness is equal to the current awareness level.

It is evident from Figure 6 that the proposed and N-A models both do a remarkably good job of predicting awareness in a real market with just a single explanatory variable. The proposed model performs somewhat better than the N-A model in predicting awareness in the holdout sample. Specifically, the average R^2 for the proposed model is 73.9% and that for the N-A model is 68.4%.

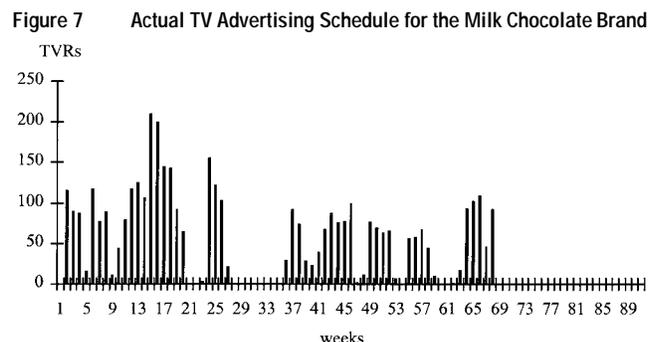
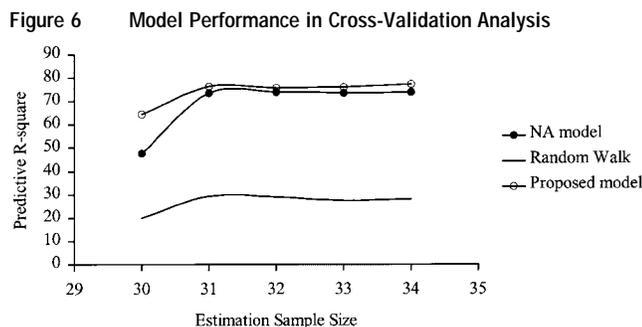
Before we consider the managerial implications of

the two models for planning media schedules, we briefly note that, in practice, field experiments need to be conducted to resolve the model selection issue.

Awareness Consequences. From the viewpoint of planning media schedules, the N-A model has one major shortcoming, namely, it cannot discriminate between different types of pulsing schedules. That is, irrespective of how a given advertising budget is expended over the planning horizon, whether in a blitz, a few pulses, or evenly over the horizon, the N-A model always predicts the same total awareness (see Tables 6 and 8 in the next section). Thus, although the N-A model fits the data as well as the proposed model, it is of little use to a media planner who wants to evaluate the awareness consequences of different pulsing schedules. In contrast, the proposed model yields different levels of total awareness for different patterns of pulsing schedules. This is because the proposed model has two state variables (recall Hartl's theorem)—awareness and advertising quality—and the dynamics of ad quality as induced by wearout and restoration effects permit discrimination between the effectiveness of different pulsing schedules. Thus, the proposed model is better suited for planning pulsing media schedules. This is demonstrated in § 5. Before doing so, however, we now summarize the model estimation results based on the milk chocolate brand data.

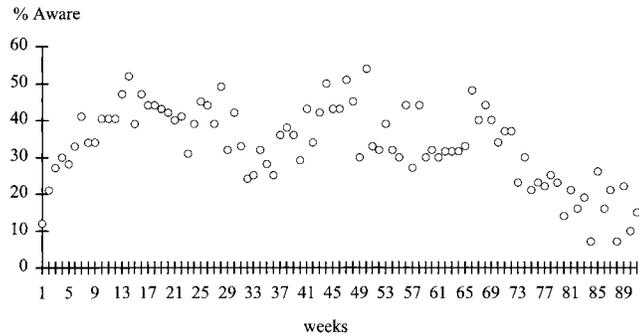
4.2. Milk Chocolate Brand Advertising Case Study

This case involves an advertising campaign over 91 weeks. The ad schedule employed is shown in Figure 7 and the corresponding advertising-awareness scores are depicted in Figure 8. These data were also collected in England in the late 1970s by Millward Brown, using



⁸We thank the editor and area editor for suggesting this analysis.

Figure 8 Percentage of Respondents Aware of the Milk Chocolate Brand's TV Advertising



the same method as in the cereal brand study. The number of weekly surveys averaged 70 households.

4.2.1. MILK CHOCOLATE BRAND DATA-BASED MODEL ESTIMATION RESULTS. Utilizing the Kalman filter approach, we repeated all the analyses described in the previous section. Once again, we found the proposed model with a linear response specification and the N-A model fit the data much better than the four other comparison models. Further, as before, we determined that the proposed and N-A models' fits to the data and predictive performances were about equally good. This replication of previous results with a different data set enhances our confidence in the validity of the proposed model. To conserve space, we report only the proposed and N-A full-sample estimation results in Table 5. (Details of all model comparisons and diagnostic tests are available from the authors.)

Comparing the proposed and N-A models, we find the quantity $-2(\Delta LL^*) = 17.22$. This exceeds the critical $\chi^2_{3df, \alpha=0.05} = 7.82$ and therefore we conclude that the goodness of fit of the proposed model is better than the N-A model's fit to the milk chocolate brand data. Next, turning to the estimated parameter values, we now see that the repetition wearout parameter, \hat{w} , of the proposed model is significant, but the copy wearout parameter \hat{c} is not. This empirical finding is different from what we found in the case of the cereal brand's advertising (where copy wearout was significant and repetition wearout was not). Further, the estimate of the forgetting rate parameter $\hat{\delta}$ is significant

and comparable for both the proposed and N-A models as are the measurement noise $\hat{\sigma}_\epsilon$ parameter estimates. Based on these estimated parameter values, the ad quality evolution according to the proposed model is depicted in Figure 9.

In the next section, we demonstrate the use and implications of the estimated models for planning media schedules in the cereal brand and milk chocolate brand case studies.

5. Planning Pulsing Media Schedules: Model-Based Solutions

Typically, media schedule planning is required on an annual basis so that contracts for media time and space can be negotiated well in advance. In this section, we first formulate the general problem of planning the best multi-pulse media schedule given a fixed total media budget, and then offer an algorithm to solve this problem. This solution procedure is illustrated in the cases of the cereal and milk chocolate brands' advertising, utilizing the estimation results from the previous section.

5.1. General Formulation of the Media Schedule Planning Problem

Similar to Mahajan and Muller (1986), we take the media planner's problem as one of determining the best pulsing media schedule that maximizes the total awareness generated over the planning horizon including the value of the terminal awareness level. Let B denote the total gross rating points (GRPs) that the fixed dollar media budget can buy for the given planning period of T weeks, and let I_t indicate whether advertising is on ($I_t = 1$) or off ($I_t = 0$) in a given week t . We assume that whenever advertising is on it is at a fixed level, which can be determined by

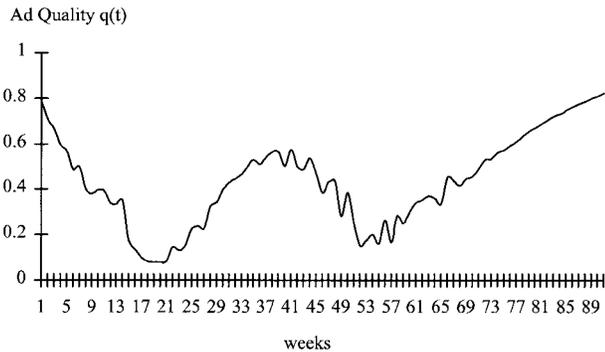
$$u(t) = \bar{u} = \frac{B}{\sum_{t=1}^T I_t}.$$

When advertising is off, the spending rate $u(t) = 0$.

Table 5 Estimates for the Proposed and N-A Models Using Milk Chocolate Brand Data

Parameters	Our Model		N-A Model	
	Estimate	t Value	Estimate	t Value
Copy wearout, c	4.8×10^{-13}	1.1×10^{-8}	—	—
Repetition wearout, w	0.0097	2.52	—	—
Forgetting rate, δ	0.0539	9.79	0.0492	4.05
Ad effectiveness, β	—	—	0.3224	4.44
Initial awareness, A_0	12.81	2.52	19.36	3.66
Initial ad quality, q_0	0.78	2.22	—	—
Measurement noise, σ_ϵ	5.1345	13.64	5.2702	11.36
Transition noise (awareness), σ_{v1}	6.74×10^{-7}	1.23×10^{-6}	2.19	4.70
Transition noise (ad quality), σ_{v2}	0.071	3.15	—	—
Maximized log-likelihood (LL)*	-205.09		-213.68	
Numbers of parameters	8		5	

Figure 9 Time-Varying Effectiveness of the Milk Chocolate Brand's Advertising



Thus, the planning problem can be conceptualized as one of finding the sequence of ones and zeros $\{I_t^*\}$ that achieves the largest total awareness.

More formally, we write the planning problem as:

$$\text{Maximize}_{\{I_t\}} J = \sum_{t=1}^{T-1} \hat{Y}_t + \frac{\hat{Y}_T}{\delta} \quad (13)$$

subject to the dynamics of awareness A_t and advertising quality q_t given by (11). In Equation (13) the performance index J is the total expected awareness, and $\hat{Y}_t = E[Y_t | \mathcal{T}_{t-1}]$ is the weekly expected awareness, which is the mean of the random variable $Y_t | \mathcal{T}_{t-1} \sim N(\hat{Y}_t, f_t)$. The planning problem is completely specified

by Equations (9)–(13), and it constitutes a continuous state, discrete time, stochastic, dynamic, optimization problem. The solution to this problem involves searching over a set of admissible pulsing sequences of $\{I_t\}_{t=1}^T$ to find that one sequence that yields the largest value of J in Equation (13).

5.2. An Algorithm for Determining the Best Pulsing Schedule

In each week the advertiser can choose to either advertise or not advertise; hence, there are 2^T spending patterns for T weeks. Not advertising at all for all the weeks is not admissible because it is assumed that the budget B has to be expended. Thus, there are $(2^T - 1)$ sequences that are admissible. For any given spending pattern, $\{I_t\}$, the forecast of the mean awareness path is computed for the entire planning horizon by using the estimated parameter values for the model dynamics (Equation (9)) and the covariance matrices (Equation (12)). Using the mean awareness forecasts, the total expected awareness J in Equation (13) can be computed. In this “what if” analysis, the actual awareness scores are not available during the planning stage, and hence the forecast errors in the Kalman filter recursion (see Appendix B) are zero. Thus, the total expected awareness is based on information available at the beginning of the planning stage, \mathcal{T}_0 . Below, we present the steps to take to find the best pulsing schedule:

1. Assume a certain sequence of ones and zeros of dimension $T \times 1$, $\{I_t\}_{t=1}^T$.
2. Forecast the mean awareness path $\{\hat{Y}_t\}_{t=1}^T$ using
 - the Kalman filter recursions given in Appendix B, Equation (B2),
 - the specified dynamics in Equation (9),
 - the stochastic characteristics from Equation (12), and
 - parameter estimates obtained from the estimation results, e.g., Tables 1 or 5.
3. Compute the performance index, J , in Equation (13) by using the above forecast.
4. Do steps 1–3 for other possible sequences of $\{I_t\}$.
5. Select the sequence corresponding to the largest value of J .

Genetic Algorithm. Now, in the case of a planning horizon of 52 weeks, each possible schedule can be considered to be a 52-bit string of ones and zeros that indicates whether advertising is on or off, respectively. There are $2^{52} \approx 4,500$ trillion such pulsing schedules. Since finding the optimal schedule is a stupendous task, if not impossible, we use the genetic algorithm (GA) (see, e.g., Goldberg 1989) as a tool to efficiently identify a set of best pulsing schedules in practice.⁹ GA has been previously used to address a marketing problem by Balakrishnan and Jacob (1996), who describe the details of how GA performs the search. We note here briefly that GA uses the evolutionary notions of natural selection, survival of the fittest, and mutation to identify the region of the global maximum for any function. Basically, any GA program (e.g., *Genitor*, see Whitley 1993) starts with a set of candidate schedules and through a process of selection, crossover, and mutation it produces better offspring (schedules) in the subsequent generations, and it eventually converges on a set that is most likely to contain the optimal schedule. The conceptual algorithm given above uses the

⁹Because of the inability to search all 4,500 trillion or more possibilities, one cannot find the unique optimal solution to the general, unconstrained problem. In practice, therefore, our recommended GA-based procedure can assist in searching over an immense set of possibilities and coming up with a set of a specified number of “best-worst” schedules that can guide decision making. However, if the set of possibilities is constrained to a finite smaller set of alternatives by management, our basic model-based algorithm can in fact exhaustively evaluate all options and determine the optimal plan.

Kalman filter to determine the “fitness” (i.e., total expected awareness given by Equation (13)) of any specific candidate schedule being considered by the GA. Thus, the Kalman filter-based algorithm is nested within the GA. Next we apply this GA-KF algorithm to find sets of “best” and “worst” annual (52-week) media schedules in the cereal brand and milk chocolate brand case studies.

5.3. Best-Worst Annual Pulsing Media Schedules

5.3.1. CEREAL BRAND CASE STUDY. The advertising budget B in this analysis is assumed to be equal to 52 times the average weekly GRPs over the 75 weeks of advertising by this brand; thus, $B = (1,051/75) \times 52 = 728.7$. Table 6 displays some best and worst weekly pulsing media schedules for the cereal brand, determined by applying the above nested GA-KF algorithm, where 1 (0) indicates that advertising should be “on” (“off”). The total expected awareness, J , associated with each of the schedules reported in Table 6 is determined by evaluating the objective functional in Equation (13).

Table 6 indicates that the best advertising schedules involve pulsing, and in fact uniform spending of the budget over the year is the worst schedule for this cereal brand in terms of the total expected awareness criterion. Further, note that under the estimated Nerlove-Arrow model, all schedules produce the same total expected awareness of 1,894.25 units, whatever their pattern of “on” or “off” weekly spending of the same advertising budget may be. Thus, as we contended in the previous section, the N-A model is unable to discriminate between different schedules and, therefore, is not useful for media schedule planning purposes. It follows from these results that there is some merit in the cereal brand company’s use of a pulsing pattern of spending (see Figure 2).

But, how good was this actual TV advertising schedule? To answer this question, the estimated model was applied to evaluate the total expected awareness outcome of the actual spending plan over the observed 75-week horizon. The performance of the actual schedule using a total budget of $B = 1,051$ GRPs is given by

$$J = \sum_{t=1}^{74} A_t + \frac{A_{75}}{0.043} = 2,662.1 \text{ units.}$$

Table 6 Some Best-Worst 52-Week Pulsing Schedules for the Cereal Brand Advertising

Number	Pulsing Media Plans	Our Model $J(I_i)$	N-A Model $J(I_i)$
1	0000 1001 1010 0001 1000 1000 0010 1101 0000 0000 0100 1110 0101	4,072.03	1,894.25
2	0000 1010 0000 1100 0010 1100 0011 0100 0101 1000 0011 0101 0001	4,060.52	1,894.25
3	0000 0001 1010 0110 0001 0000 0001 0001 1110 0100 0101 0101 0011	4,054.80	1,894.25
4	0001 0010 0000 1000 1101 1000 1100 0001 1010 0001 0000 1011 0110	4,042.25	1,894.25
5	0001 0100 0001 1001 1000 1001 0001 0110 0101 0000 0001 0110 0100	4,037.20	1,894.25
6	1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 0100 0000	996.51	1,894.25
7	1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1101	989.28	1,894.25
8	1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 0000	977.50	1,894.25
9	1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1000	976.63	1,894.25
10	1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1110	974.99	1,894.25
Even	1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111	974.22	1,894.25

In contrast, with the same budget for the 75-week horizon, three selected plans from the best set of 50 plans found with the proposed GA-KF algorithm, along with their total expected awareness, are shown in Table 7.

Thus, we have found a number of different pulsing schedules of spending the same budget that could have been used by the cereal brand company to generate more than twice the total expected awareness resulting from the schedule actually employed. However, we note that this pulsing schedule depends on the model assumptions and brand-specific parameter estimates. Next, we summarize the advertising spending schedule evaluation results in the milk chocolate brand case study.

5.3.2. MILK CHOCOLATE BRAND CASE STUDY. In this case, a total 4,041 GRPs were expended over a 91-week period, implying a 52-week budget of $B = 2,309.1$ GRPs. Table 8 presents the sets of some “best-worst” advertising schedules, which were determined

by using the GA-KF algorithm and the parameter estimates in Table 5.

Once again, we see that the best schedules involve pulsing. However, unlike the case of the cereal brand advertising, the even spending schedule is now not the worst schedule because its total expected awareness of 1,565.6 units is superior to that of a number of different pulsing schedules. Further, we see that, under Nerlove and Arrow’s model, the total expected awareness outcome is 1,888.75 units regardless of the pattern of spending, thus demonstrating the inability of the N-A model to assist in choosing between different media schedules.

Lastly, the total expected awareness outcome of the actual 91-week schedule with a budget of 4,041 GRPs employed by the chocolate brand company is evaluated to be $J = 3,254.3$. We note that the model-based 91-week spending plans using the same budget yield much superior results (Table 9).

Table 7 Three Selected Plans Found with the GA-KF Algorithm for the Cereal Brand Advertising

Plan	$J(I_i)$	Pulsing Media Schedule for the Cereal Brand, I_i
A	6,116.53	0000 0000 0000 0001 1100 0000 0001 0001 1000 1000 0110 0001 0001 0000 0100 1110 0100 0000 101
B	5,975.57	0000 0000 0100 0000 0000 0000 1001 0001 1000 1000 0110 0001 0001 1011 0000 0110 0100 0100 101
C	5,938.37	0000 0110 0000 1001 0001 0001 1001 0001 1000 1000 0110 0001 0001 0000 0001 0100 0010 0000 101

Table 8 Some Best-Worst 52-Week Pulsing Schedules for the Milk Chocolate Brand

Nos.	Pulsing Media Plans	Our Model J(I _j)	N-A Model J(I _j)
1	0010 0110 0011 0110 0110 0001 0001 0011 0011 0101 0101 0111 1110	4,159.92	1,888.75
2	0110 0101 1100 0110 0011 1110 0011 0000 0010 0111 1001 1100 0111	4,083.52	1,888.75
3	0010 1001 0101 1111 0100 0010 0111 0001 0010 0010 1000 1100 0001	4,073.63	1,888.75
4	0100 0010 0111 1011 1011 1111 1111 0000 0111 0100 0001 0101 1101	4,050.49	1,888.75
5	0111 0001 0111 1001 0101 0101 1001 0111 1111 0101 0101 1010 0010	4,032.61	1,888.75
Even	1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111	1,565.59	1,888.75
7	0000 0000 0011 0101 0000 0110 0111 1100 1111 0111 1010 1100 0000	786.32	1,888.75
8	0000 0100 1001 1000 0101 1111 1111 1111 1010 1001 1101 1111 1111	669.55	1,888.75
9	0000 0110 0000 1111 1111 1011 1111 1111 1111 1111 1111 1111 0111	518.57	1,888.75
10	0000 0000 0111 1110 1111 1111 1111 1111 1111 1110 1111 1110 0110	507.86	1,888.75
11	0000 0000 0101 1111 1111 1111 1111 1111 1111 1111 1111 1111 1110	427.49	1,888.75

Table 9 Three Selected Plans Found with GA-KF Algorithm for the Milk Chocolate Brand Advertising

Plan	J(I _j)	Pulsing Media Schedule for Milk Chocolate Brand Advertising, I _j
A	8,215.75	1001 0001 0001 0101 1100 0000 0100 0010 0101 0001 1010 0001 0100 1000 0000 1110 0100 0000 0101 0001 0000 1000 001
B	8,048.64	1001 0001 0001 0101 1100 0000 0100 0010 0101 0001 1011 0000 0110 0100 0000 1110 0100 0000 0101 0001 0000 1000 001
C	7,823.38	1001 0001 0001 0101 1100 0000 0100 0010 0101 0001 1011 0000 0110 0100 0000 0110 0001 0001 1000 0100 0000 0000 111

Again, we see that this milk chocolate brand company could have achieved much better results following one of these model-based schedules of spending the same advertising budget.

6. Conclusions

The proposed model suggests that pulsing practices may be correct in some instances, and our results suggest that pulsing schedules would have been more effective than even ones for the two cases studied. This is because the effectiveness of advertising varies over time and with different levels of repetition. We have shown that when effectiveness of advertising declines with continuous advertising (due to wearout effects) and restores during a media hiatus (due to the forgetting effect), pulsing media schedules can be superior to spending the advertising budget evenly over the planning horizon. However, if wearout effects are negligible, then the effectiveness of advertising remains

constant over time and the continuous spending strategy is optimal, consistent with the extant literature (e.g., Sasieni 1989).

Our potential contribution to the practice of media scheduling is that we offer an implementable model and algorithm that allows a media planner to determine the best schedule, whether it is even or pulsing. For example, there are myriad possible annual 52-week pulsing schedules and our model suggests which might be most effective. To use this model, the media planner would need to estimate parameters like copy and repetition wearout rates and the forgetting rate for a specific brand's advertising campaign. For estimation purposes, the advertiser will have to track the performance of the existing media plan in terms of its ability to generate brand awareness. Such tracking is already routinely done by many firms (e.g., Zielske 1986; Rossiter and Percy 1997, pp. 585–610).

Although our model's goal is to select the optimal

schedule, we do not want to appear so naive as to expect experienced media planners to mindlessly follow the decisions indicated by the model. However, the media planner's expertise could be combined with the suggestions of our model similar to the use of Little and Lodish's (1969) heuristic model for media resource allocation. Our model has the potential to improve media scheduling decisions, even if it is used as an evaluation tool rather than as an optimizing model. For example, in the two case studies described, the model-based solutions obtained from the proposed GA-KF algorithm indicate that there was considerable room for improvement in the performance of the advertising schedules selected by the decision makers.

Two conclusions are clear about the modeling of the effects of advertising repetition and different schedules. First, the optimal advertising spending schedule is highly sensitive to different model assumptions and parameter estimates. We feel that our assumptions are justified from the research literature as well as empirical data, but recognize that others may disagree with our interpretation and specification. In our tests of different specifications, such as whether the specifications of ad spending should be linear or nonlinear, our model was superior to other forms. However, we have not tested an exhaustive set of alternatives, and a search of the "right" specification needs to be continued in future research.

The second conclusion about modeling advertising repetition and scheduling is that we have not been able to reject the N-A model, based on the statistical evidence. While our model does better on most criteria than N-A, the differences are not sufficient to reject N-A. Clearly, our model is much more amenable to aiding decisions about pulsing versus continuous scheduling, since the N-A model invariably recommends the even spending schedule. While further data sets may allow a more complete statistical comparison, we agree with the view that the only way to "settle" the issue of which model is better on statistical grounds is to experiment with different schedules of ad repetition and timing and then compare the performance of the two models in predicting the results of different campaigns to different groups. Fortunately, there is an established split cable television technology to conduct field experiments by advertisers (e.g., Lodish et al. 1995).

Several issues related to media scheduling were not considered in this paper. While these are not trivial issues, better data, as well as better models, are required to address them in any substantial manner. In particular, four unresolved factors stand out that require more research: First, our model does not address the important issue of *advertising competition* (e.g., Park and Hahn 1991; Villas-Boas 1993), which needs investigation in the presence of dynamic advertising quality (i.e., ad wearout and restoration effects). With better data that also include, if possible, actual or estimated GRPs of competitors' advertising, future modeling may be able to deal with this problem. Second, we also do not deal with *multiple copy campaigns* that might be running simultaneously. That is why the two media campaigns we selected to test our model each included only one advertising copy. Advertising-awareness data, such as proven ad recall (which is gathered in most advertising tracking studies but often not quantitatively analyzed), could disentangle the effects of different copy and allow for different copy to wear out at different rates. However, such modeling will be difficult since the carryover effects of different past advertising campaigns can adversely affect awareness of the current campaign (Sutherland 1993, pp. 183–194). Third, we do not explore the gains due to pulsing schedules with unequal intensities and related issues of how much advertising frequency is required for maintenance periods in the schedule when advertising is low but not zero. A related issue involves concerns about seasonality. Our model could probably address this latter issue by first deciding how much to budget for different times of the year (seasons) and then using those seasons as the planning horizon for our model.

A fourth issue concerns the measure of advertising effectiveness to be analyzed and used as a criterion in our model. We use advertising awareness as the dependent variable in model estimations. Critics may prefer different measures such as brand image, brand attitude, or even sales (see, for example, Zielske 1986). We would love to have data to estimate our model on other measures of advertising effectiveness. As Little (1986, p. 107) stated in his discussion of Mahajan and Muller's (1986) model of ad awareness and Sasieni's (1971) model of brand sales, "the mathematics does not

care" what measure the model includes. However, different measures differ in their responsiveness to advertising and multiple ad exposures over time (e.g., Ray 1982, p. 399). Although we believe that ad awareness is a measure that tracks the extent to which wear-out will occur since the resulting inattention should be first indicated by ad awareness, other measures of effectiveness may correlate more highly with sales (e.g., Axelrod 1980). However, ad awareness has been shown to correlate with sales in tracking studies (e.g., Colman and Brown 1983). Despite these shortcomings, we believe that our model is a distinct improvement over extant scheduling models and, even in its present form, can be a helpful aid to advertising decision makers. We hope to be able to extend our model to deal with at least some of these real-world difficulties. Since research efforts to extend advertising scheduling models in these directions should improve the theory and practice of media planning, we invite other researchers to join us in these efforts.¹⁰

Appendix A

Given a blitz schedule as defined by Equation (8), the solution to model Equations (2)–(7) is:

$$A_1(t) = \frac{g(\bar{u})}{\delta - a(\bar{u})} (e^{-a(\bar{u})t} - e^{-\delta t}) + A_0 e^{-\delta t}, \quad t \in [0, l], \quad (\text{A1})$$

$$A_2(t) = A_1(l)e^{-\delta(t-l)}, \quad t \in [l, T], \quad (\text{A2})$$

where $A_1(t)$ expresses how awareness evolves over the first l weeks when advertising is on; $A_2(t)$ is how awareness decays from its level at the end of l weeks, i.e., the point at which the budget is exhausted and advertising is stopped; and A_0 is the initial awareness level at the beginning of the campaign. Next, following Mahajan and Muller (1986), we find the total (cumulative) awareness, $J(l)$, generated by the specified blitz policy, which is given by the area under the distribution of awareness levels over time. That is,

¹⁰We gratefully acknowledge that this paper has benefited from comments received from the area editor and two anonymous reviewers at *Marketing Science*, Fred Feinberg, Eitan Gerstner, Mike Hagerty, Jonathan Hamilton, John Lynch, Steven Shugan, Joffre Swait, Chih-Ling Tsai, Miguel Villas-Boas, Bart Weitz, Russ Winer, and participants at the 1997 U.C. Berkeley seminar. We thank Prof. Jeff Harrison for providing the data sets, and Prof. David Woodruff for helping us with the genetic algorithm. The authors are solely responsible for any remaining errors.

$$\begin{aligned} J(l) &= \int_0^\infty A(t)dt \\ &= \int_0^l A_1(t)dt + \int_l^T A_2(t)dt + \int_T^\infty A_2(t)dt. \end{aligned} \quad (\text{A3})$$

The sum of first two right-hand side (RHS) terms of Equation (A3) represents total awareness generated by the blitz campaign over the planning horizon of length T and the third RHS term denotes the "salvage value" of the terminal awareness level $A(T)$, i.e., the additional total awareness over period from T through infinity. Then, substituting Equations (A1) and (A2) in (A3) and integrating, we get:

$$J(l) = \frac{g(\bar{u})}{\delta a(\bar{u})} (1 - e^{-a(\bar{u})l}) + \frac{A_0}{\delta}. \quad (\text{A4})$$

Note that by letting $l = T$ in Equation (A4), i.e., having a "blitz" that lasts for the entire planning horizon, we obtain the total awareness under an even schedule.

Appendix B: Moments of the Conditional Density $f(Y_t | \mathcal{F}_{t-1})$

This appendix provides the moments of the conditional density $f(Y_t | \mathcal{F}_{t-1})$. We recall that the observation equation is $Y_t = z\alpha_t + \epsilon_t$, the transition equation is $\alpha_t = T_t\alpha_{t-1} + c_t + \nu_t$, and error terms are distributed as $\epsilon_t \sim N(0, H)$ and $\nu_t \sim N(0, Q)$. Since all error terms are distributed normally and the transition and observation equations are linear in the state variables (A_t, q_t) , the random variable $Y_t | \mathcal{F}_{t-1}$ is normally distributed. This follows from the fact that "the sum of normal random variables is a normal."

To obtain the moments of the normal random variable $Y_t | \mathcal{F}_{t-1}$, let \hat{Y}_t denote the mean and f_t be the variance of $Y_t | \mathcal{F}_{t-1}$. The mean is obtained by taking the expectation of observation equation as follows:

$$\begin{aligned} \hat{Y}_t &= E[Y_t | \mathcal{F}_{t-1}] \\ &= E[z\alpha_t + \epsilon_t | \mathcal{F}_{t-1}] \\ &= zE[\alpha_t | \mathcal{F}_{t-1}] + 0 \\ &= za_{t|t-1}, \end{aligned} \quad (\text{B1})$$

where $a_{t|t-1}$ is the mean of the state variable α_t given \mathcal{F}_{t-1} (to be determined). Similarly, the variance of $Y_t | \mathcal{F}_{t-1}$ is derived as follows:

$$\begin{aligned} f_t &= \text{Var}[Y_t | \mathcal{F}_{t-1}] \\ &= \text{Var}[z\alpha_t + \epsilon_t | \mathcal{F}_{t-1}] \\ &= z \text{Var}[\alpha_t | \mathcal{F}_{t-1}]z' + \text{Var}[\epsilon_t | \mathcal{F}_{t-1}] \\ &= zP_{t|t-1}z' + H, \end{aligned} \quad (\text{B2})$$

where $P_{t|t-1}$ is the covariance matrix of state variable α_t given \mathcal{F}_{t-1} (to be determined). We note that $\text{Var}(\epsilon_t | \mathcal{F}_{t-1}) = \text{Var}(\epsilon_t) = H$ follows from Equation (12).

The mean vector and covariance matrix of the state variable $\alpha_t =$

$[A_t, q_t]'$ are obtained recursively from the following expressions in (B3), which are known as Kalman's recursions (see Harvey 1994 for details of the derivation):

$$\begin{aligned} a_{t|t-1} &= T_t a_{t-1} + c_t, P_{t|t-1} = T_t P_{t-1} T_t' + Q, \\ a_t &= a_{t|t-1} + K_t (Y_t - z a_{t|t-1}), \\ P_t &= P_{t|t-1} - K_t z P_{t|t-1}, K_t = P_{t|t-1} z' f_t^{-1}. \end{aligned} \quad (B3)$$

To obtain $a_{t|t-1}$ and $P_{t|t-1}$ for each t , $t = 1, \dots, T$, we iterate through the above recursion by assuming a diffused initial prior on $\alpha_0 \sim \text{MVN}(a_0, P_0)$. For example, given (a_0, P_0) , we get $(a_{1|0}, P_{1|0})$ and thus (a_1, P_1) ; now given (a_1, P_1) , we get $(a_{2|1}, P_{2|1})$ and thus (a_2, P_2) ; and so on. Knowing $a_{t|t-1}$ and $P_{t|t-1}$ for each t , the moments of $Y_t | \mathcal{F}_{t-1}$ can then be determined by using Equations (B1) and (B2). The unknown initial mean vector, a_0 , is estimated by treating it as hyperparameters in the likelihood function; see the estimates for \hat{A}_0 and \hat{q}_0 in Tables 1 and 5.

References

- Agresti, Alan. 1990. *Categorical Data Analysis*. Wiley Interscience, New York.
- Appel, Valentine. 1971. On advertising wearout. *J. Advertising Res.* 1 (February) 1–13.
- Arnold, Stephen J., Tae H. Oum, Bohumir Pazderka, Douglas W. Snetsinger. 1987. Advertising quality in sales response models. *J. Marketing Res.* 24 (February) 106–113.
- Assmus, Gert, John Farley, Donald Lehmann. 1984. How advertising affects sales: Meta-analysis of econometric results. *J. Marketing Res.* 21 (February) 65–74.
- Axelrod, Joel N. 1980. Advertising wearout. *J. Advertising Res.* 20 (October) 13–18.
- Balakrishnan, P. V., Varghese S. Jacob. 1996. Genetic algorithms for product design. *Management Sci.* 42 (August) 1105–1117.
- Belch, George E. 1982. The effects of television commercial repetition on cognitive response and message acceptance. *J. Consumer Res.* 9 (June) 56–65.
- Berlyne, D. E. 1970. Novelty, complexity, and hedonic value. *Perception and Psychophysics* 8 279–286.
- Blackburn, J. D., K. J. Clancy. 1982. LITMUS: A new product planning model. Andris A. Zoltners, ed. *Marketing Planning Models*. North-Holland Publishing Company, New York. 43–62.
- Blair, Margaret H. 1988. An empirical investigation of advertising wearin and wearout. *J. Advertising Res.* 27 (December–January) 45–50.
- Blair, Margaret H., Karl E. Rosenberg. 1994. Convergent findings increase our understanding of how advertising works. *J. Advertising Res.* 34 (May–June) 35–45.
- Blattberg, Robert, J. Golanty. 1978. Tracker: An early test market forecasting and diagnostic model for new product planning. *J. Marketing Res.* 15 (May) 192–202.
- Blattberg, Robert C., Abel P. Jeuland. 1981. A micromodeling approach to investigate the advertising-sales relationship. *Management Sci.* 27 (September) 988–1005.
- Brown, Gordon. 1985. Tracking studies and sales effects: A U.K. perspective. *J. Advertising Res.* 25 (February–March) 52–65.
- Cacioppo, John T, Richard E. Petty. 1979. Effects of message repetition and position on cognitive response, recall, and persuasion. *J. Personality and Social Psychology* 37 (January) 97–109.
- Calder, Bobby, Brian Sternthal. 1980. Television commercial wearout: An information processing view. *J. Marketing Res.* 17 (May) 173–186.
- Calantone, Roger, Alan G. Sawyer. 1978. The stability of benefit segments. *J. Marketing Res.* 15 (August) 395–404.
- Colman, Stephen, Gordon Brown. 1983. Advertising tracking studies and sales effects. *J. Market Res.* Society 25 (April) 165–183.
- Corkindale, David, John Newall. 1978. *Advertising Threshold and Wearout*. MCB Publications, Bradford, England.
- Craig, Samuel, Brian Sternthal, Clark Leavitt. 1976. Advertising wearout: An experimental analysis. *J. Marketing Res.* 13 (November) 15–22.
- Feichtinger, Gustav, Richard F. Hartl, Suresh P. Sethi. 1994. Dynamic optimal control models in advertising: Recent developments. *Management Sci.* 40 (February) 195–226.
- Feinberg, Fred. 1992. Pulsing policies for aggregate advertising models. *Marketing Sci.* 11 221–234.
- Goldberg, David E. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA.
- Gould, John P. 1970. Diffusion processes and optimal advertising policy. E. S. Phelps et al., eds., *Microeconomic Foundations of Employment and Inflation Theory*. W. W. Norton, New York.
- Grass, Robert, W. H. Wallace. 1969. Satiation effects of TV commercials. *J. Advertising Res.* 9 (September) 3–8.
- Greenberg, Allan, Charles Suttoni. 1973. Television commercial wearout. *J. Advertising Res.* 13 (October) 47–54.
- Greyser, Stephen A. 1973. Irritation in advertising. *J. Advertising Res.* 13 (February) 3–10.
- Groenhaug, Kjell, Olaf Kvitastein, Sigmund Gronmo. 1991. Factors moderating advertising effectiveness as reflected in 333 tested advertisements. *J. Advertising Res.* 31 (October–November) 42–50.
- Haley, Russell L. 1978. Sales effects of media weight. *J. Advertising Res.* 18 (June) 9–18.
- Hamilton, James D. 1994. *Time Series Analysis*. Princeton University Press, Princeton, N.J.
- Hartl, Richard F. 1987. A simple proof of the monotonicity of the state trajectories in autonomous control problems. *J. Econom. Theory* 41 211–215.
- Harvey, Andrew C. 1994. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, New York.
- Hughes, G. David. 1992. Real-time response measures re-define advertising wearout. *J. Advertising Res.* 32 (May–June) 61–77.
- Jones, John Philip. 1995. *When Ads Work: New Proof That Advertising Triggers Sales*. Lexington Books, New York.
- Katz, William A. 1980. A sliding schedule of advertising weight. *J. Advertising Res.* 20 (August) 39–44.

- Krugman, Herbert E. 1972. Why three exposures may be enough. *J. Advertising Res.* **12** (December) 11–14.
- Little, John D. C. 1975. Brandaid: A marketing mix model, part 1—Structure. *Operations Res.* **23** 628–655.
- Little, John D. C. 1979. Aggregate advertising models: The state of the art. *Operations Res.* **27** (July–August) 629–667.
- Little, John D. C. 1986. Comments on advertising pulsing policies for generating awareness for new products. *Marketing Sci.* **5** (Spring) 107–108.
- Little, John D. C., Leonard M. Lodish. 1969. A media planning calculus. *Operations Res.* **17** (January–February) 1–35.
- Lodish, Leonard M, Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, Mary E. Stevens. 1995. How T.V. advertising works: A meta-analysis of 389 real world split cable T.V. advertising experiments. *J. Marketing Res.* **32** (May) 125–139.
- Luhmer, Alfred, Alois Steindl, Gustav Feichtinger, Richard Hartl, Gerhard Sorger. 1988. ADPULS in continuous time. *European J. Operational Res.* **34** (March) 171–177.
- Mahajan, Vijay, Eitan Muller. 1986. Advertising pulsing policies for generating awareness for new products. *Marketing Sci.* **5** 2 89–111.
- Mahajan, Vijay, Eitan Muller, Subash Sharma. 1984. An empirical comparison of awareness forecasting models of new product introduction. *Marketing Sci.* **3** (Summer) 179–206.
- Mesak, Hani. 1992. An aggregate advertising pulsing model with wearout effects. *Marketing Sci.* **11** 3 310–326.
- Naples, Michael J. 1979. *Effective Frequency: The Relationship Between Frequency and Advertising Effectiveness*. National Association of Advertisers, Inc, New York.
- Nerlove, Marc, Kenneth Arrow. 1962. Optimal advertising policy under dynamic conditions. *Economica* **29** (May) 129–142.
- Park, Sehoon, Minhi Hahn. 1991. Pulsing in a discrete model of advertising competition. *J. Marketing Res.* **28** (November) 397–405.
- Parsons, Leonard, Randall Schultz. 1976. *Marketing Models and Economic Research*, North-Holland, New York.
- Pechmann, Cornelia, David Stewart. 1990. Advertising repetition: A critical review of wearin and wearout. MSI Report, No. 90–106. Cambridge, MA.
- Pekelman, Dov, Suresh Sethi. 1978. Advertising budgeting, wearout and copy replacement. *J. Operations Res. Society* **29** 7 651–659.
- Rao, Ambar G., Peter B. Miller. 1975. Advertising/sales response functions. *J. Advertising Res.* **15** (April) 7–15.
- Rao, Ram C. 1986. Estimating continuous time advertising-sales models. *Marketing Sci.* **5** (Spring) 125–142.
- Ray, Michael L. 1982. *Advertising and Communication Management*. Prentice-Hall, Englewood Cliffs, NJ.
- Ray, Michael L., Alan G. Sawyer. 1971. Behavioral measurement for marketing models: Estimating the effects of advertising repetition for media planning. *Management Sci.* **18** (December) 73–89.
- Rossiter, John R., Larry Percy. 1997. *Advertising Communications and Promotion Management*. McGraw-Hill, New York.
- Sasieni, Maurice. 1971. Optimal advertising expenditure. *Management Sci.* **18** (December) 64–72.
- Sasieni, Maurice. 1989. Optimal advertising strategies. *Marketing Sci.* **8** (December) 358–370.
- Sawyer, Alan G. 1981. Repetition, cognitive response, and persuasion. R. E. Petty, T. M. Ostrom, T. C. Brock, eds. *Cognitive Responses to Persuasion*. Erlbaum New York. 237–262.
- Simon, Hermann. 1982. ADPULS: An advertising model with wearout and pulsation. *J. Marketing Res.* **19** (August) 352–363.
- Simon, Julian L. 1979. What do Zielske's real data really show about pulsing? *J. Marketing Res.* **16** (August) 415–420.
- Simon, Julian L., Johan Arndt. 1980. The shape of the advertising response function. *J. Advertising Res.* **20** (August) 11–28.
- Strong, Edward C. 1974. The use of field experimental observations in estimating advertising recall. *J. Marketing Res.* **11** (November) 369–378.
- Strong, Edward C. 1977. The spacing and timing of advertising. *J. Advertising Res.* **17** (December) 25–31.
- Sutherland, Max. 1993. *Advertising and the Mind of the Consumer: What Works, What Doesn't and Why*. Allen and Urwin, Sydney.
- Vidale, M. L., H. B. Wolfe. 1957. An operations-research study of sales response to advertising. *Oper. Res.* **5** (June) 370–381.
- Villas-Boas, Miguel. 1993. Predicting advertising pulsing policies in oligopoly: A model and empirical test. *Marketing Sci.* **12** 1 88–102.
- Webb, Peter H., Michael L. Ray. 1979. Effects of TV clutter. *J. Advertising Res.* **19** (June) 7–14.
- Weilbacher, W. M. 1970. What happens to advertisements when they grow up? *Public Opinion Quarterly* **34** (Summer) 216–223.
- West, Mike, Jeff Harrison. 1997. *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York.
- Whitley, L. Darrell. 1993. *Foundations of Genetic Algorithms*. M. Kaufmann Publishers, San Mateo, Calif.
- Zielske, Hubert A. 1959. The remembering and forgetting of advertising. *J. Marketing* **23** (January) 239–243.
- Zielske, H. A. 1986. Comments on "Advertising pulsing policies for generating awareness for new products." *Marketing Sci.* **5** (Spring) 89–111.
- Zielske, Hubert, Walter Henry. 1980. Remembering and forgetting of television ads. *J. Advertising Res.* **20** (April) 7–13.

This paper was received May 30, 1997, and has been with the authors 8 months for 1 revision; processed by David C. Schmittlein.