

# The Anatomy of the Advertising Budget Decision: How Analytics and Heuristics Drive Sales Performance

Journal of Marketing Research  
 2020, Vol. 57(3) 468-488  
 © American Marketing Association 2020  
 Article reuse guidelines:  
[sagepub.com/journals-permissions](http://sagepub.com/journals-permissions)  
 DOI: 10.1177/0022243720907578  
[journals.sagepub.com/home/mrj](http://journals.sagepub.com/home/mrj)



Ceren Kolsarici , Demetrios Vakratsas, and Prasad A. Naik

## Abstract

The authors investigate how heuristics and analytics contribute to the advertising budget decision by decomposing it into four components: (1) baseline spending, (2) adaptive experimentation, (3) advertising-to-sales ratio, and (4) competitive parity. They propose a methodology to estimate and infer the weights of these four components. Applying this methodology to sales and advertising data across eight brands from three categories substantiates for the first time, and uniformly across all brands, that managers depart from optimality through adaptive experimentation, which is in line with dual control theory that suggests they do so to learn about advertising effectiveness. The adaptive experimentation finding, combined with evidence on the use of heuristic methods, suggests that budget decision making is characterized by bounded rationality. Furthermore, budgeting decisions are brand-specific, reflecting the considerations of a brand's market position and performance. Finally, simulation studies show that brands from categories with high uncertainty in advertising effectiveness can benefit from double-digit revenue lifts by placing higher emphasis on adaptive experimentation.

## Keywords

adaptive dual control, adaptive experimentation, advertising budget, bounded rationality, heuristics

The advertising process starts with the budgeting decision and ends with the distribution of content to consumers through various media outlets. Although the effects of media spending on sales have received considerable attention in the literature, not much is known about how managers decide to set their advertising budgets and how their decisions affect sales performance. The limited available evidence suggests that sales performance varies depending on the budgeting method, particularly in markets characterized by uncertainty and dynamic advertising response (Wang and Zhang 2008). Furthermore, to counterbalance the predominance of heuristics rules (Danenberget al. 2016), the advent of analytics and the proliferation of data call for the adoption of more evidence-based budgeting strategies.

What is known about the budgeting process, mostly at the industry level but not for individual brands, is that managers do not follow a single approach to set their budgets but rather use a combination of them. This practice is referred to as “multiple stakes in the ground” (Farris and West 2007). Such approaches include (1) heuristics drawing upon managerial judgment, such as percentage of past or expected sales (also known as advertising-to-sales [A/S] ratio) and competitive

benchmarking, and (2) analytical approaches such as profit-maximizing models (e.g., West, Ford, and Farris 2014). In other words, decision making regarding budget setting invariably combines managerial judgment and model-driven insights (Rosenzweig 2014). However, interviews with top-level as well as brand managers reveal that identifying the different methods involved in the advertising budgeting process and quantifying the extent to which each method is used is a difficult task (e.g., Farris and West 2007; Low and Mohr 1999; Piercy 1987). For example, Farris and West (2007, p.316) note, “Exactly how the recommendations of various approaches are (should be) combined to determine a single dollar figure is, unfortunately, not clear.” Furthermore, the advertising budget decision process is the outcome of negotiations between top-level and marketing (or brand) management, involving considerations such as power and political influence (e.g., Piercy 1987). Consequently, disentangling the contribution of each

Ceren Kolsarici is Associate Professor of Marketing, Queen's University, Canada (email: [ceren.kolsarici@queensu.ca](mailto:ceren.kolsarici@queensu.ca)). Demetrios Vakratsas is Associate Professor of Marketing, McGill University, Canada (email: [demetrios.vakratsas@mcgill.ca](mailto:demetrios.vakratsas@mcgill.ca)). Prasad A. Naik is Professor of Marketing, University of California Davis, USA (email: [panaik@ucdavis.edu](mailto:panaik@ucdavis.edu)).

method used in the budget decision-making process is a challenging task.

This discussion points to the lack of a systematic representation of the advertising budgeting process and its impact on sales performance. Such a representation would offer an unbiased assessment of the relative contribution of analytics and heuristics methods to the overall budgeting decision and provide a diagnostic tool to evaluate their effects on sales performance. For example, marketing managers may admit less to the use of heuristic-based methods and more to analytics-based approaches for fear of appearing less sophisticated. As West, Ford, and Farris (2014, p. 12) remark, “Companies need to become more open about their budgeting processes and acknowledge the role that heuristics play.”

To realize these benefits, we propose a dynamic model of advertising spending over time, without presuming optimality of managers’ behavior, and then link it to sales performance. In other words, our modeling framework accommodates bounded rationality in managerial decision making by acknowledging that managerial behavior goes beyond profit maximization, allowing multiple objectives such as maximizing revenues subject to satisficing profit levels, pursuing sales growth, or warding off competition (Baumol 1962; Gunny 2005; March and Simon 1958). In striving to satisfice these goals, managers often rely on heuristics and depart from optimality. Thus, in our framework, nonoptimality does not imply lack of rationality.

Our intended contribution is fourfold. First, we formulate the combined use of heuristics and analytics-based budgeting rules in advertising decision making consistent with the “multiple stakes” practice paradigm. This formulation constitutes a novel approach in the advertising budgeting literature. Second, we account for the role of managers’ ambiguity regarding the response function parameters (e.g., Tull et al. 1986) by explicitly incorporating the uncertainty of ad effectiveness in setting advertising budgets. This parametric uncertainty marks a departure from standard profit-maximizing models that assume full knowledge of advertising effectiveness. Third, we allow the advertising budget to evolve over time due to changing budgeting methods, market conditions, and temporal variation in advertising effectiveness. Fourth, we empirically show, for the first time in the literature, that managers depart from optimality by spending proportionately to the uncertainty in ad effectiveness. This knowledge sheds light on the rationalization of departing from profit-maximizing optimality. In other words, suboptimal behavior does not imply non-rational behavior.

We apply the proposed model to sales, advertising, and marketing mix data for eight brands across durable (hybrid cars) and nondurable (beer and yogurt) categories at different stages of their product life cycle. Using the proposed methodology, this study provides empirical evidence, uniformly across eight brands from three categories, that managers depart from optimality by spending proportional to the uncertainty in ad effectiveness, a hitherto untested hypothesis in marketing grounded in the dual control theory (see Filatov and Unbehauen

2004). Furthermore, managers also depart from optimality by using both advertising-to-sales ratio and competitive parity heuristics. Most importantly, the empirical results confirm the use of the “multiple stakes in the ground” practice of making ad budgeting decisions on the basis of market data. We emphasize that the extant literature (e.g., Farris and West 2007) employs managerial surveys rather than real-world market data and, consequently, the reported results are frequency counts summarized at the industry level—not at the disaggregated brand level—based on managers’ opinions about the budgeting methods they use. In contrast, the proposed framework provides the empirical weights for a specific brand on the basis of its actual ad spending and sales.

## Conceptual Development

We consider both survey-based descriptive studies and normative studies to identify the salient methods managers use to set the advertising budget. Our review suggests that budgeting methods can be grouped into the following four broad categories: (1) baseline advertising spending, representing a certainty-based method; (2) adaptive experimentation, representing an uncertainty-based method; (3) percentage of sales (A/S ratios); and (4) competitive parity. The first two represent analytics-driven approaches, and the latter two are commonly used heuristics. We next review each method by discussing relevant literature.

### Baseline Advertising Spending

A method focused exclusively on baseline spending reflects analytical thinking about cost control and profit maximization (Wang and Zhang 2008). For example, studies on the topic of optimal advertising spending have suggested that optimizing advertising allocation depends on the advertising effectiveness, margin, and discount rate (for a review, see Vakratsas and Naik 2007). More recently, Raman et al. (2012) derived the optimal solution when advertising effectiveness, margin, and discount rate vary over time. The existence of a closed-form solution suggests that managers can use optimal spending as a method for setting advertising budgets. At the same time, quantitative methods are difficult to implement, and so managers set aside a fixed amount as the baseline spending. Therefore, the baseline spending method reflects the optimal advertising level conditional on managerial knowledge of advertising effectiveness. In terms of performance implications, Wang and Zhang (2008) have suggested that constant spending can sustain sales in a stable market and overcome sales decline.

Of course, a priori managerial knowledge with complete certainty regarding advertising effectiveness, as in Raman et al. (2012), is a strong assumption given the presence of many different intervening factors such as competition, advertising wear-out, and unforeseen exogenous shocks. Therefore, managers may deviate their spending with the goal of learning about ad effectiveness through adaptive experimentation.

### Adaptive Experimentation

Managers occasionally undertake field experiments by intentionally manipulating advertising levels to inject exogenous variation with the hope of discovering advertising effectiveness and its bounds (e.g., Ackoff and Emshoff 1975; Eastlack and Rao 1989). Random assignment of advertising timing, campaigns, or regions help rule out confounding factors to permit causal inference (e.g., Sahni 2015). The main drawback of field experiments, however, is that managers learn about ad effectiveness only at the intervention time, or for specific campaigns and test regions, and relative to the predetermined controlled campaigns/regions. In other words, the discovered ad effectiveness is localized in the space-time setting of the field experiments. Over time, this experimental knowledge is implemented to all regions, as it would be unprofitable and certainly unwise to restrict some regions to being permanent control units given a successful intervention or to keep permuting various regions as control units and not reap the benefits of a successful test.

So, as experiments end naturally, business-as-usual settles in and managers continue to face the need to understand how ad effectiveness varies over time for a host of factors beyond those tested or potentially tested under experimental conditions, such as copy and restoration wear-out (Naik, Mantrala, and Sawyer 1998), product life cycle stage (Kolsarici and Vakratsas 2015; Osinga, Leeflang, and Wieringa 2010), ad content (Kolsarici and Vakratsas 2010), competitive interference (Danaher, Bonfrer, and Dhar 2008), or product harm crisis (Rubel, Naik, and Srinivasan 2011; Van Heerde, Helsen, and Dekimpe 2007). In addition, temporal variation may be due to seasonality and coordination with promotional campaigns. The presence of time-varying advertising effectiveness suggests that managers must learn continuously about the impact of their advertising campaigns.

To this end, adaptive experimentation facilitates managerial learning regarding changes in advertising effectiveness. In their discussion of adaptive experimentation, Pekelman and Tse (1980, p.322) observe, "A useful experiment is the one which is conducted *continuously* over time." An "experiment" does not have to be an intentional design localized in space-time; rather, it can involve introducing a variation into a constant setting. The fact that managers change ad spending over time and it induces changes in brand sales affords them steady opportunities to learn about ad effectiveness. However, this variation in ad spending need not be exogenous, and so managers should account for the endogenous nature of ad spending as in this study (see Equations 5 and 7) or control for endogeneity as in Varian's (2016) review (e.g., instrumental methods, regression discontinuity, differences-in-differences).

Another advantage of adaptive experimentation is that managers can learn not only about time-varying ad effectiveness but also time-varying uncertainty in ad effectiveness. Then, according to the hypothesis grounded in the dual control theory (see Filatov and Unbehauen 2004), managers can set the ad budget proportional to the uncertainty in ad effectiveness (we elaborate on this in the section "Analytics-Driven Budgeting

Decision: Dual Control"). Finally, experimentation differs from the practice of pulsing; the latter reflects the timing and spacing of media schedules (see Vakratsas and Naik 2007), whereas the former is focused on the setting of ad budgets proportionately to uncertainty in ad effectiveness.

### Percentage of Sales Heuristic (A/S Ratios)

Interviews with marketing managers suggest that almost no advertising budget decision starts from a "zero basis" (Low and Mohr 1999). Instead, managers typically rely on their judgment and beliefs regarding ad effectiveness, as well as recent performance, to form a starting point for budgeting deliberation. The surveys by Piercy (1987) and Bigné (1995) indicate that managers rely extensively on A/S ratios to determine their budgets. Although largely considered an "unsophisticated" method applied in "top-down" budgeting processes in which top management's influence is greater, A/S is anchored on performance expressed by either past sales or projected future sales. Nerlove and Arrow (1962) showed that the A/S ratio is proportional to the ratio of advertising and price elasticities, a finding that Raman (1990) extended for the stochastic case. Thus, the A/S is a "responsive" budgeting rule (Wang and Zhang 2008), albeit benchmarked only against a brand's own performance and effectiveness without taking into account the competitive environment.

The performance implications of the A/S method depend on the market conditions. Implementation of the A/S method in a growing market will lead to sustained growth but will lead to sales erosion in a declining market. The predominance of this top-down budget process is possibly due to the lack of top-level management confidence in the role of advertising. It also reflects a belief in reverse causality in which sales determine advertising rather than the other way around (Danenberg et al. 2016).

### Competitive Parity Heuristic

There are two approaches to the competitive parity method: one uses absolute benchmarks by setting the budget roughly equal to that of the competition, and the other uses relative metrics by setting the share-of-voice (SOV) to be roughly equal to share of market (SOM) (Farris and West 2007). The latter is also known as the "SOV equals SOM" method and is preferred by managers operating in mature markets (Danaher, Bonfrer, and Dhar 2008). Naik, Prasad, and Sethi (2008) analyzed the N-brand dynamic game and showed that the Nash equilibrium spending for each brand is proportional to the total market share of other brands. In other words, a large (small) brand should spend less (more) than its fair share in dynamic oligopoly markets. This normative finding counters the SOV equals SOM method. Therefore, the use of the competitive parity heuristic (i.e., ad spending proportional to market share) constitutes evidence for managerial departure from optimal behavior. Another drawback of the competitive parity heuristic is that it may lead to escalation and, consequently, erode ad

effectiveness due to increased clutter as brands try to continuously keep up with their competitors (Danaher, Bonfrer, and Dhar 2008).

Given the well-documented practice of “multiple stakes in the ground,” a model of advertising budget decision process should allow for the possible use of a combination of the four salient methods discussed in this section. The next section shows how to incorporate the four budgeting methods in a dynamic model of ad spending, assess their relative contributions, and estimate and infer time-varying ad effectiveness and its uncertainty.

## Model Development

Our model incorporates the previously discussed two main features of the ad budget decision process: (1) use of a combination of multiple methods and (2) the role of ad effectiveness uncertainty. To accommodate the first feature, we assume that managerial decision making is characterized by bounded rationality and is driven by both analytics and heuristics considerations. The analytics-driven component is informed by dual control theory, whereas the heuristics-driven component consists of the A/S ratio and competitive parity methods. For the second feature, we model sales and ad effectiveness in a state space framework that captures both time-varying ad effectiveness and its associated uncertainty. We first discuss the analytics-driven component of the ad budget decision and then formulate the state space model.

### *Analytics-Driven Budgeting Decision: Dual Control*

Prior literature on optimal ad budgeting assumes that managers know with certainty the parameters of the advertising response function (e.g., Naik, Raman, and Winer 2005; Tull et al. 1986). Under this assumption, managers can determine the optimal budget using standard optimal control theory. By contrast, in the presence of uncertainty in ad effectiveness, managers must dually focus on both brand profits and parameter learning. Adaptive dual control theory applies perfectly to this dual objective setting, as it envisions a manager who finds the control policy (i.e., ad spending over time) and learns about the uncertain model parameters simultaneously. Although the original formulation of the optimal adaptive dual control problem, due to Feldbaum (1965), is unsolvable, the bicriterial synthesis method, developed a decade later (see Wittenmark 1975), successfully addresses the problem. As the name “dual control” suggests, the bicriterial approach maximizes two objective functions that correspond to the two goals of finding the control policy and learning about the uncertain parameters of a dynamic system (see Filatov and Unbehauen 2004 for more details). As discussed in the “Conceptual Development” section, Little (1966) and Pekelman and Tse (1980) applied the ideas in dual control theory to promotions and advertising, respectively, to lay the pioneering groundwork for adaptive experimentation in marketing.

To make the ideas of dual control theory precise, we specify a commonly used dynamic advertising–sales model with sales carryover and other covariates:

$$S_t = \lambda S_{t-1} + \beta_t u_t + Z'_{1t} \eta + \varepsilon_{1t}, \quad (1)$$

where  $S_t$  denotes sales in period  $t$ ,  $u_t$  is the ad spending,  $Z_{1t}$  is a vector of covariates such as price, promotions, coupons, seasonality, and other controls to be discussed later (see the Data Description section),  $(\lambda, \beta_t, \eta)'$  are the parameters for sales carryover, ad effectiveness, and the vector of coefficients for control variables, respectively, and  $\varepsilon_{1t}$  represents the normally distributed error term to capture the impact of factors not explicitly included in the model. Furthermore, ad effectiveness evolves through a random walk with a drift,

$$\beta_{t+1} = \beta_t + Z'_{2t} \kappa + \varepsilon_{2t}, \quad (2)$$

where the drift equals  $Z'_{2t} \kappa$ , with  $Z_{2t}$  containing the covariates and  $\varepsilon_{2t}$  representing the normally distributed error term.

The Appendix derives the profit-maximizing advertising trajectory,  $u_t^*$ , in the presence of uncertain ad effectiveness using stochastic control theory. Both the derivation and the resulting solution are novel to the extant literature, which ignores uncertainty in time-varying ad effectiveness (see, e.g., Raman et al. 2012). In dual control theory, however, the manager perturbs the dynamic model to learn about the uncertain ad effectiveness, thus departing from profit optimality. The dual control question can then be framed as follows: How much profit reduction should a manager tolerate as they depart from optimality to increase parametric learning? Let  $\Delta$  denote the ad budget set for adaptive experimentation to learn about uncertain ad effectiveness. According to the dual control theory (Filatov and Unbehauen 2004, pp. 37–40), the greater the parametric uncertainty, the larger the budget required for learning about the parameter values. The admissible range for bicriterial optimization is  $(\tilde{u}_b, \tilde{u}_b + \Delta)$ , where  $\tilde{u}_b$  is the baseline spending and  $\Delta$  represents the budget for experimentation. Following Filatov and Unbehauen (2004, pp. 37–40), the experimental budget is set at  $\Delta = \tilde{\gamma}_1 \sigma_{\beta_t}^2$ , where  $\tilde{\gamma}_1$  is a proportionality constant and  $\sigma_{\beta_t}^2$  measures the uncertainty in ad effectiveness, both to be estimated empirically. In other words, as uncertainty in ad effectiveness increases substantially (i.e., large  $\sigma_{\beta_t}^2$ ), the experimental budget should increase correspondingly (i.e., larger  $\Delta$ ). Thus, the total analytically driven budget decision is given by

$$u_t^a = \tilde{u}_b + \Delta = \tilde{u}_b + \tilde{\gamma}_1 \sigma_{\beta_t}^2. \quad (3)$$

Equation 3 combines the baseline spending and the increment caused by the uncertainty in ad effectiveness, thus suggesting that managers can rationally depart from optimality. In other words, nonoptimal does not imply nonrational. Besides the rational departure from optimality, the realities of ad budget setting suggest that managers may need to combine analytically driven decisions with heuristics, which represent managerial experience and firm-related practices. Therefore, we augment

the analytics component by incorporating the role of heuristics in advertising budget decisions.

**Augmented Budget Decision Model: Incorporating Heuristics**

If managers were to rely exclusively on analytical approaches to set ad spending, Equation 3 would accurately reflect the budget decision. However, as our literature review suggests, managers also use heuristic approaches. Thus, we augment Equation 3 by incorporating the heuristic influences of the A/S ratio and competitive benchmarking. According to the A/S ratio method, ad spending is proportional to expected sales,  $E_t$ , based on past sales performance (see Raman 1990). We operationalize expected sales using historical data by setting  $E_t = \frac{1}{H_S} \sum_{h=1}^{H_S} S_{t-h}$ , where we empirically calculated the duration  $H_S$  to be 11 months on the basis of the results of a grid search. To operationalize competitive benchmarking, we use the SOV equals SOM heuristic (see Danaher, Bonfrer, and Dhar 2008; Danaher and Rust 1994). Specifically, we set competitively benchmarked spending as  $C_t = X_{t-1}^I MS_{t-1}$ , where  $X_{t-1}^I$  and  $MS_{t-1}$  denote the lagged values of industry spending and the market share of the focal brand, respectively. Thus, the heuristic component of the budget is given by

$$u_t^h = \tilde{\gamma}_2 E_t + \tilde{\gamma}_3 C_t. \tag{4}$$

Finally, we combine the analytics and heuristics budget decision components (Equations 3 and 4) with relevant covariates to arrive at the augmented advertising budget decision

$$u_t = u_b + \gamma_1 \sigma_{\beta_t}^2 + \gamma_2 E_t + \gamma_3 C_t + Z_{3t}' \delta, \tag{5}$$

where  $u_t$  represents the monthly advertising budget;  $(\gamma_1, \gamma_2, \gamma_3)'$  capture the effects of adaptive experimentation, advertising-to-sales heuristic, and competitive benchmarking, respectively;  $u_b$  represents the baseline spending; and the vector  $\delta$  denotes the effects of various control variables in  $Z_{3t}$ . The different notation, with and without tilde, for the intercepts and slopes reflects that their corresponding magnitudes can be unequal when estimating the combined equation (Equation 5) in the presence of covariates. The various covariates in  $Z_{3t}$  include seasonality, annual controls, promotions, price, and other marketing efforts, as well as a pulsing dummy accounting for excessive spending on advertising, defined by the indicator variable  $I_t = \begin{cases} 1 & \text{if } X_t > \mu_x + 3\sigma_x \\ 0 & \text{otherwise} \end{cases}$ , where  $(\mu_x, \sigma_x, X_t)$  are respectively the mean, standard deviation, and the observed level of ad spending by the focal brand.

Equation 5 suggests that the advertising budget depends on analytical and heuristic decision rules as well as various brand and market related variables. Thus, it reflects a bounded rationality approach to resource allocation in which, in addition to analytics, managers rely on their own judgment and experience (Low and Mohr 2000). More specifically, managers set aside a baseline budget  $u_b$  and complement it with an allocation to

experimentation, which is proportional to the uncertainty in ad effectiveness  $\sigma_{\beta_t}^2$ . The ad budget increases by  $\gamma_1$  dollars for a unit increase in the uncertainty of ad effectiveness;  $\gamma_2$  captures the impact of the advertising-to-sales heuristic; and  $\gamma_3$  corresponds to the effect of the competitive parity heuristic. From a substantive contribution perspective, a positive  $\gamma_1$  would provide evidence that brands spend proportionately to the uncertainty in ad effectiveness, a hypothesis grounded in dual control theory (Filatov and Unbehauen 2004) but not hitherto verified empirically in the extant marketing literature. Such evidence, which we shall provide, not only confirms the veracity of the dual control theory but also sheds light on the rationality of departure from optimality. To simultaneously estimate Equations 1 and 5, we express them in the state space framework.

**State Space Framework**

The state space framework is commonly used in the marketing literature (e.g., Ataman, Van Heerde, and Mela 2010; Bruce 2008; Kolarici and Vakratsas 2010; Kolarici and Vakratsas 2015). It is represented by two sets of equations: the transition equation and the observation equation. In a vector-matrix form, the transition equation can be expressed as follows:

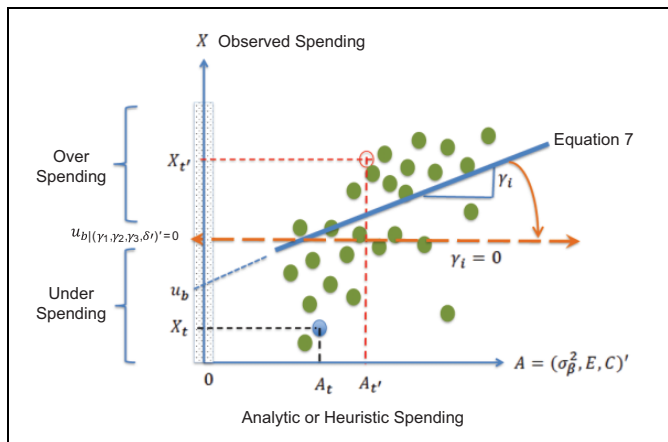
$$\underbrace{\begin{bmatrix} S_t \\ \beta_{t+1} \end{bmatrix}}_{f_t} = \underbrace{\begin{bmatrix} \lambda & u_t \\ 0 & 1 \end{bmatrix}}_{T_t} \underbrace{\begin{bmatrix} S_{t-1} \\ \beta_t \end{bmatrix}}_{f_{t-1}} + \underbrace{\begin{bmatrix} d_{1t} \\ d_{2t} \end{bmatrix}}_{D_{1t}} + \underbrace{\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2,t+1} \end{bmatrix}}_{\varepsilon_t} \tag{6}$$

where  $D_{1t} = (d_{1t}, d_{2t})'$  is the drift vector consisting of  $d_{1t} = Z_{1t}' \eta$  and  $d_{2t} = Z_{2t}' \kappa$ , and  $(\varepsilon_{1t}, \varepsilon_{2,t+1})' = \varepsilon_t \sim N(0, Q)$  represent the normal error terms with  $Q = \sigma_\varepsilon^2 I_{2 \times 2}$ . Equation 6 constitutes the canonical transition equation  $f_t = T_t f_{t-1} + D_{1t} + \varepsilon_t$ , where  $f_t$  denotes the state vector at instant  $t$  and  $T_t$  is the time-varying transition matrix.

We next link the state vector to the observation equation  $y_t = Hf_t + D_{2t} + v_t$ , as follows:

$$\underbrace{\begin{bmatrix} Y_t \\ X_t \end{bmatrix}}_{y_t} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_H \underbrace{\begin{bmatrix} S_t \\ \beta_{t+1} \end{bmatrix}}_{f_t} + \underbrace{\begin{bmatrix} 0 \\ u_t \end{bmatrix}}_{D_{2t}} + \underbrace{\begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}}_{v_t} \tag{7}$$

where the vector  $y_t = (Y_t, X_t)'$  contains the observed sales  $Y_t$  and observed advertising spending  $X_t$ ,  $D_{2t} = (0, u_t)'$  is the drift vector with  $u_t$  defined in Equation 5, and  $(v_{1t}, v_{2t})' = v_t \sim N(0, R)$  are the normal error terms with  $R = \sigma_v^2 I_{2 \times 2}$ . The observed sales  $Y_t$  serve as one of the two dependent variables, whereas  $S_t$  represents the model-based forecast  $S_t = E[Y_t] = \hat{Y}_t$  via Equation 7. The observed ad spending  $X_t$  serves as the other dependent variable, so the second row in Equation 7 directly incorporates the endogenous nature of advertising rather than correcting for it as in Varian (2016). Together, Equations 6 and 7 comprise the canonical state space form.



**Figure 1.** Graphical illustration of the second row in Equation 7.

Finally, we explain how the proposed modeling framework accommodates over- or underspending. Figure 1 presents the second row of Equation 7 that connects to Equation 5 substantively, and Table 1 describes the notations. The y-axis in Figure 1 depicts the observed spending ( $X_t, X_{t'}$ ) at different instants ( $t, t'$ ). The x-axis in Figure 1 denotes any one of the three elements of experimental or heuristic spending  $A_t = (\sigma_{\beta_t}^2, E_t, C_t)'$  with their corresponding effects represented by  $(\gamma_1, \gamma_2, \gamma_3)'$ . (The other covariates in Equation 5 can be interpreted analogously to  $A_t$ .) The intercept  $u_b$  yields the baseline spending, which is obtained by projecting the regression line to the y-axis. When  $(\gamma_1, \gamma_2, \gamma_3, \delta')' = 0$ , the experimentation, heuristics, and covariate effects in Equation 5 are absent, the regression line in Figure 1 becomes horizontal, and the resulting intercept  $u_{b(\gamma_1, \gamma_2, \gamma_3, \delta')' = 0}$  represents the theoretical optimum. The observed spending  $X_t$  depends not only on the theoretical optimum  $u_{b(\gamma_1, \gamma_2, \gamma_3, \delta')' = 0}$  but also on the informational shocks  $v_{2t}$  in Equation 7. These shocks represent myriad events in the market environment as well as within the company, such as tweets by the president that affect the industry, the company's earnings management needs, top executives' PR scandals, product harm crises, and supply chain disruptions such as a major warehouse fire, to mention a few. Consequently, when managers react and respond to negative (positive) informational shocks, they tend to depart from the theoretical optimum, resulting in under (over) advertising. By contrast, when  $(\gamma_1, \gamma_2, \gamma_3, \delta')' \neq 0$ , the estimated intercept  $u_b$  represents the baseline spending and not the theoretical optimum  $u_{b(\gamma_1, \gamma_2, \gamma_3, \delta')' = 0}$ , as depicted in Figure 1. Even when managers systematically spend above the baseline due to nonzero effects of heuristics and adaptive experimentation, the negative informational shocks can lead to cutbacks in spending below the baseline. Thus, Equation 7 accommodates both over- and underspending depending on the estimated parameters and corresponding variable levels. Figure 1 illustrates over- and underspending via two realizations,  $(X_t, X_{t'})$ .

In summary, the proposed modeling framework addresses five key issues. First, the transition equation (Equation 6) simultaneously captures the dynamics in sales and advertising effectiveness over time. Second, the ad budget is decomposed into four components: baseline spending, adaptive experimentation proportional to ad effectiveness uncertainty, and the two heuristic methods of A/S ratio and competitive parity. Third, the advertising budget decision depends on covariates such as seasonality, price, promotions, coupons, and annual controls. Fourth, the state space framework directly incorporates the endogenous nature of advertising. Finally, as discussed previously, Equation 7 accommodates over and underspending.

### Estimation Approach

The transition and observation equations are estimated using the Kalman filter, which is uniquely the best estimator (i.e., it yields the minimum mean squared errors) of the state vector  $f_t$  according to the sequentially observed history of the market data  $y_t$  in the information set  $\mathcal{I}_t = \{y_1, y_2, \dots, y_t\} = \mathcal{I}_{t-1} \cup \{y_t\}$ , for  $t = 1, 2, \dots, N$ . The filter consists of two steps: time update and measurement update. To briefly elaborate, the stochastic dynamics of the state vector  $f_t$  are given by Equation 6. At each instant  $t$ , the time update step yields the prior mean vector ( $\hat{f}_{t|t-1}$ ) and prior covariance matrix ( $P_{t|t-1}$ ) of the state vector  $f_t$ . Then, upon observing the market data  $y_t$ , the measurement update step yields the posterior mean ( $\hat{f}_{t|t}$ ) and posterior covariance ( $P_{t|t}$ ), as shown in Table 2.

The extent of uncertainty in ad effectiveness at each instant,  $\{\hat{\sigma}_{\beta_t}^2; t = 1, \dots, N\}$ , is given by the last diagonal element of the prior covariance matrix. More generally, using the prior state mean and its covariance from the Kalman filter, we evaluate the log-likelihood of observing the market data  $y_t = (Y_t, X_t)'$  via  $LL(\theta) = \sum_{t=1}^N \text{Ln}(p(y_t | \mathcal{I}_{t-1}); \theta)$ , where  $p(y_t | \mathcal{I}_{t-1}; \mu_t, \Sigma_t)$  is the multivariate normal probability of observing  $y_t$  conditional on the information set  $\mathcal{I}_{t-1}$ , with mean  $\mu_t = H\hat{f}_{t|t-1} + D_{2t}$  and covariance matrix  $\Sigma_t = HP_{t|t-1}H' + R$ , and  $\theta = (\beta_0, u_b, \gamma_1, \gamma_2, \gamma_3, \eta', \kappa', \delta', \phi, \sigma_v^2, \sigma_e^2)'$  is the  $K \times 1$  parameter vector with  $\beta_0$  as the starting point of the ad effectiveness trajectory  $\beta_t$  in Equation 2. Some parameters such as the carryover effect need to be restricted to the positive unit interval, so we transform  $\lambda = g(\phi) = \frac{\exp(\phi)}{1 + \exp(\phi)}$  and estimate the unrestricted  $\phi \in (-\infty, \infty)$ . Similarly to this interval constraint, positivity constraints  $(u_b, \gamma_1, \gamma_2, \gamma_3)' > 0$  can be imposed, if necessary, by transforming  $u_b$  or any  $\gamma_i = \exp(\tilde{\phi})$  and estimating the unrestricted  $\tilde{\phi} \in (-\infty, \infty)$ . As shown subsequently, we can construct confidence intervals using the Delta method (Davidson and MacKinnon 2004, p. 205).

**Table 1.** Summary of Notation.

Bounded Rationality Framework		Normative Decision Framework	
$Y_t$	Observed sales at time $t$ .	$u_t$	Feasible ad spending trajectory over the span $t \in [0, \infty)$
$S_t$	Model-based sales at time $t$	$c(x) = x^2$	Quadratic function to ensure global maximum
$(X_t, X'_t)$	Observed ad spending at different instants $(t, t')$	$J(u_t)$	Net present value of $u_t$
$u_a$	Total analytic spending	$u_t^*$	Theoretical optimal spending
$u_h$	Total heuristic spending	$u_{b(\gamma_1, \gamma_2, \gamma_3, \delta')=0}$	Intercept in Equation 5 when $(\gamma_1, \gamma_2, \gamma_3, \delta') = 0$
$E_t$	Expected sales at $t$		
$C_t$	Competitor spends at $t$		
$Z_{1t}, Z_{2t}, Z_{3t}$	Covariates at $t$		
$\mu_x$	Sample mean of $X_t$		
$\sigma_x^2$	Sample variance of $X_t$		
Parameters to Be Estimated			
$u_b$	Baseline spending when $(\gamma_1, \gamma_2, \gamma_3, \delta') \neq 0$		
$\gamma_1, \gamma_2, \gamma_3$	Impact of analytic or heuristic spending on $X_t$		
$\beta_t$	Ad effectiveness at $t$		
$\sigma_{\beta_t}^2$	Uncertainty in ad effectiveness at $t$		
$\lambda$	Carryover effect		
$\eta, \kappa, \delta$	Effects of covariates		
$\sigma_\epsilon, \sigma_v, Q, R$	Variances in various equations		

**Table 2.** Kalman Filter Recursions.

	Time Update (before observing $y_t$ )	Measurement Update (after observing $y_t$ )
Mean state vector	$\hat{f}_{t t-1} = T_t \hat{f}_{t-1 t-1} + D_{1t}$	$\hat{f}_{t t} = \hat{f}_{t t-1} + K_t (y_t - H \hat{f}_{t t-1} - D_{2t})$
State covariance matrix	$P_{t t-1} = T_t P_{t-1 t-1} T'_t + Q$	$P_{t t} = (I - K_t H) P_{t t-1}$
Kalman gain factor		$K_t = P_{t t-1} H' (R + H P_{t t-1} H')^{-1}$

Next, we maximize the log-likelihood function to obtain the parameter estimates:

$$\hat{\theta} = \text{ArgMax } LL(\theta). \tag{8}$$

To infer significance of the estimated parameters, we obtain the standard errors from the square root of the diagonal elements of the inverse of the information matrix:

$$M = \left[ -\frac{\partial^2 LL(\theta)}{\partial \theta \partial \theta'} \right]^{-1}. \tag{9}$$

To safeguard against unknown forms of misspecification, we conduct robust inference using the sandwich estimator (White 1980):

$$\text{Cov}(\hat{\theta}) = \text{MVM}, \tag{10}$$

where  $V = G'G$  and  $G$  is an  $N \times K$  matrix created by stacking the gradient of the log-likelihood for each

observation. If the model is correctly specified, then  $M^{-1} = V$  and the inferences based on (9) and (10) are identical, as they should be. The square root of the diagonal of  $\text{Cov}(\hat{\theta})$  yields the standard errors of the parameter estimates. We calculate the t-values by dividing the estimates from Equation 8 by their corresponding standard errors from Equation 10.

Finally, by applying the Delta method (see Davidson and MacKinnon 2004, p. 205), we obtain the standard error for the carryover effect as follows:

$$\text{se}(\hat{\lambda}) = \text{se}(\hat{\phi}) \times \frac{\partial g}{\partial \phi} = \text{se}(\hat{\phi}) \exp(\hat{\phi}) / (1 + \exp(\hat{\phi}))^2. \tag{11}$$

The standard errors of positively constrained parameters are given by  $\text{se}(\hat{\gamma}_i) = \text{se}(\hat{\phi}) \exp(\hat{\phi})$ .

**Table 3.** Monte Carlo Simulation for Model Identification.

	True Values	No Variation (E = 10)	Low Variation E ∈ [9, 11]	High Variation E ∈ [5, 15]
		Average Estimated Values		
Process noise, $\sigma_\epsilon$	30	27.60	28.19	29.55
Measurement noise, $\sigma_v$	30	29.77	30.44	29.91
Ad effectiveness, $\beta$	1.5	1.53	1.52	1.50
Carryover effect, $\lambda$	.65	.64	.65	.65
Baseline spending, $u_b$	50	20.79	49.16	50.28
Advertising-to-sales heuristic, $\gamma_2$	10	17.88	10.87	9.92
Determinant of the Hessian of LL*		0	2.38	54.98
Maximized log-likelihood, LL*		-488.98	-491.42	-489.60
Relative weight of baseline spending	33.23%	14.10%	33.63%	32.99%
Relative weight of A/S heuristic	66.77%	85.90%	66.37%	67.01%

### Monte Carlo Simulations

Although the proposed model seeks to assess the contributions of the different methods to the total budget, we acknowledge that disentangling the contributions when the shared variation between them is high would be difficult. To investigate this issue, we use a simplified version of Equation 5 with the baseline spending and A/S ratio components only representing analytical and heuristic decision rules, respectively. We conduct Monte Carlo simulations under three conditions of variation in the advertising-to-sales heuristic—constant expected sales (i.e., no variation), low variation in expected sales, and high variation in expected sales—operationalized using the expected sales  $E_t$  variable in Equation 5. We simulate data sets from our model using parameter values in typical ranges, perform 100 simulations for each condition, estimate parameters using the proposed Kalman filter approach 100 times, and observe the model's ability to recover the true (known) parameter values.

For the three conditions, Table 3 presents the actual and mean estimated parameters, average log-likelihood, average determinant of the Hessian matrix of the log-likelihood function, and the true and average estimated level importance of the two advertising rules in contributing to the overall ad budget (see the “Ad Budget Decomposition” section for details on “level importance”). In the no variation condition, the Hessian is singular because its determinant equals zero, so the model is not identifiable. This is also evident from the parameter estimates and the weights of baseline spending and the A/S heuristic, which depart significantly from their true values. In other words, if sales do not exhibit variation, both baseline spending and the A/S heuristic predict rather flat advertising expenditures and it is not possible to distinguish between the two rules. By contrast, for both the low and high variation conditions, Table 3 reveals that the estimated parameters and weights are retrieved satisfactorily. Parameter recovery improves as the variation in expected sales increases. Thus, the presence of variation in expected sales is sufficient for estimating the parameters. We do not expect identification issues in the empirical application because sales and method-based spending exhibit

sufficient variation over time and the estimated Hessian was nonsingular.

We also conduct simulations with all advertising decision rules. Specifically, we generate 1,000 data sets from the system of Equations 6 and 7 using random walk ad effectiveness as per Equation 2 and expected sales and competitive spending from the uniform distribution  $U(5, 15)$ . The true parameter vector is  $\theta = (\beta_0, u_b, \gamma_1, \gamma_2, \gamma_3, \lambda, \sigma_v^2, \sigma_\epsilon^2)' = (10, 1, 1, 1, 1, 0.5, 1, 1)'$ . We estimated the model parameters by applying the proposed approach. The average estimates are  $\hat{\theta} = (10.29, 0.94, 1.01, 1.03, 1.01, 0.51, 0.98, 1.01)'$ , which are close to the true values. Their corresponding t-values,  $t = (9.66, 2.76, 245.33, 24.29, 25.42, 24.50, 11.76, 12.65)'$ , are also significant across the board.

## Empirical Analysis

### Data Description

For the purpose of generalizability and understanding the nuances of advertising budgeting decisions, we test our model on eight brands from three product categories. Specifically, we use data from the categories of hybrid cars, beer, and yogurt in the U.S. market. The categories differ with respect to the product type (i.e., durable for hybrid cars and nondurable for consumer goods), maturity (i.e., new markets for hybrids and established ones for consumer goods), level of fragmentation (i.e., consolidated for hybrids, medium fragmentation for yogurt, and high for beer), and frequency of purchase. We focused on the top performing brands—namely, Toyota Prius and Honda Civic in hybrid cars, which had a combined market share of 95%, Yoplait and Danone in the yogurt category, which had a combined market share of 62%, and Bud Light, Budweiser, Coors Light, and Miller Lite in the beer category, which had a combined market share of 39% during the observation periods. These brands also represent the top advertisers in their respective categories.

The data sets consist of monthly observations of sales, advertising, and various covariates, covering a 68-month period from January 2005 to August 2010 for hybrid cars and



a 132-month period from January 2001 to December 2011 for the yogurt and beer categories. The sources for sales data are JD Power and Associates (JDPA) for hybrid cars and the IRI Marketing data set (see Bronnenberg, Kruger, and Mela 2008) for the packaged goods brands. The media spending data represent aggregated monthly spending across all media outlets for the advertised brand (e.g., TV, print, newspaper, internet, outdoor) and are sourced from Kantar media. We include several covariates to capture the drivers of dynamic state variables. Specifically, we account for the effects of price, promotions, coupons, annual controls, and seasonality on multiple outcomes such as sales, ad effectiveness, and ad spending for the beer and yogurt brands. We use federal tax credits as a proxy for new model introduction for the hybrid car data.

**Operationalization of Variables**

The IRI data consist of weekly store-level sales, promotions, coupons, and price for each SKU between 2001 and 2011. We aggregated the packaged goods data from (1) SKUs to brand, (2) weekly to monthly level, and (3) store to market level, as follows: First, we aggregated the sales, price, coupon, and promotions data from SKUs to brand and store to market level. In the aggregation process, we standardized the variables where possible (e.g., price per ounce) and averaged across SKUs weighted by the store-level SKU sales. We used a similar procedure to obtain the mean price promotion and mean coupon probability for each brand at the market level. Second, we aggregated these brand and market-level weekly variables to monthly level by taking into account the start and end dates of the observed weeks. When a week spanned over two consecutive months, we performed the aggregation by taking the weighted average of variables across the number of days/week in each month assuming uniform distribution across the days of the week. Third, we dropped feature and/or display variables due to the high correlation of their frequency with price promotions. The resulting brand-level variables exhibited adequate variance for estimating the determinants of brand-specific budgeting decisions. Finally, we adjusted prices for inflation by dividing them by the consumer price index of each month for the respective product category according to the U.S. Bureau of Labor Statistics. Table 4 presents the descriptive statistics.

Figure 2 presents the evolution of advertising and sales for each brand. Across all brands, ad spend varies over the observation period, contrary to the constant optimal advertising suggested in the literature (e.g., Naik and Raman 2003). This variation is higher for the hybrid car brands than their consumer packaged good (CPG) counterparts. Moreover, beer brands are mature and exhibit less dynamics. Yogurt, although established as a category, displays growth in sales, and the emerging hybrid cars category also exhibits growth in sales. These differences across categories can potentially translate into diverse scenarios with respect to the combination of methods used for the advertising budgeting decision, a point on which we subsequently elaborate.

**Table 4.** Descriptive Statistics.

	Bud Light		Budweiser		Coors Light		Miller Lite		Yoplait		Danone	
	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
Sales (\$ millions)	6.61	(3.85, 9.89)	3.83	(2.46, 6.36)	3.42	(2.07, 5.82)	3.50	(2.07, 5.54)	6.92	(4.08, 10.61)	6.27	(3.56, 9.41)
Mean price	19.72	(16.81, 23.39)	19.06	(16.93, 21.67)	18.83	(16.73, 21.45)	18.52	(16.44, 20.94)	2.00	(1.83, 2.14)	2.06	(1.66, 2.52)
Mean promotion probability	.18	(.13, .24)	.18	(.13, .25)	.19	(.14, .26)	.20	(.14, .25)	.26	(.07, .39)	.25	(.09, .39)
Mean coupons	.00	(.00, .01)	.00	(.00, .01)	.00	(.00, .00)	.00	(.00, .01)	.00	(.00, .09)	.00	(.00, .02)
Advertising (\$ millions)	13.28	(2.27, 40.47)	8.75	(1.51, 32.44)	10.25	(1.23, 23.36)	9.98	(1.38, 23.64)	8.62	(1.21, 20.38)	6.16	(.00, 22.05)
Market share	14.5%	(12.6%, 16.3%)	9.2%	(6.9%, 12.5%)	7.2%	(6.5%, 8.2%)	7.4%	(6.3%, 8.7%)	32.5%	(23.9%, 36.6%)	29.4%	(24%, 33.7%)
	Prius		Civic									
	Mean	Range	Mean	Range								
Sales (1,000 units)	11.05	(5.75, 22.05)	2.12	(.19, 5.43)								
Price (\$1,000)	25.28	(23.69, 27.08)	22.56	(19.46, 25.04)								
Tax credit (\$1,000)	1.18	(.00, 3.15)	.71	(.00, 1.70)								
Advertising (\$ millions)	2.78	(.00, 30.30)	1.40	(.00, 12.11)								
Market share	79.2%	(60.2%, 96.9%)	15.3%	(1.3%, 28.1%)								

Note: IRI data covers roughly 5% of the stores in the United States. Therefore, sales values reported for the CPG brands are sums across stores that are members of the IRI panel and not estimates of the total market volume.

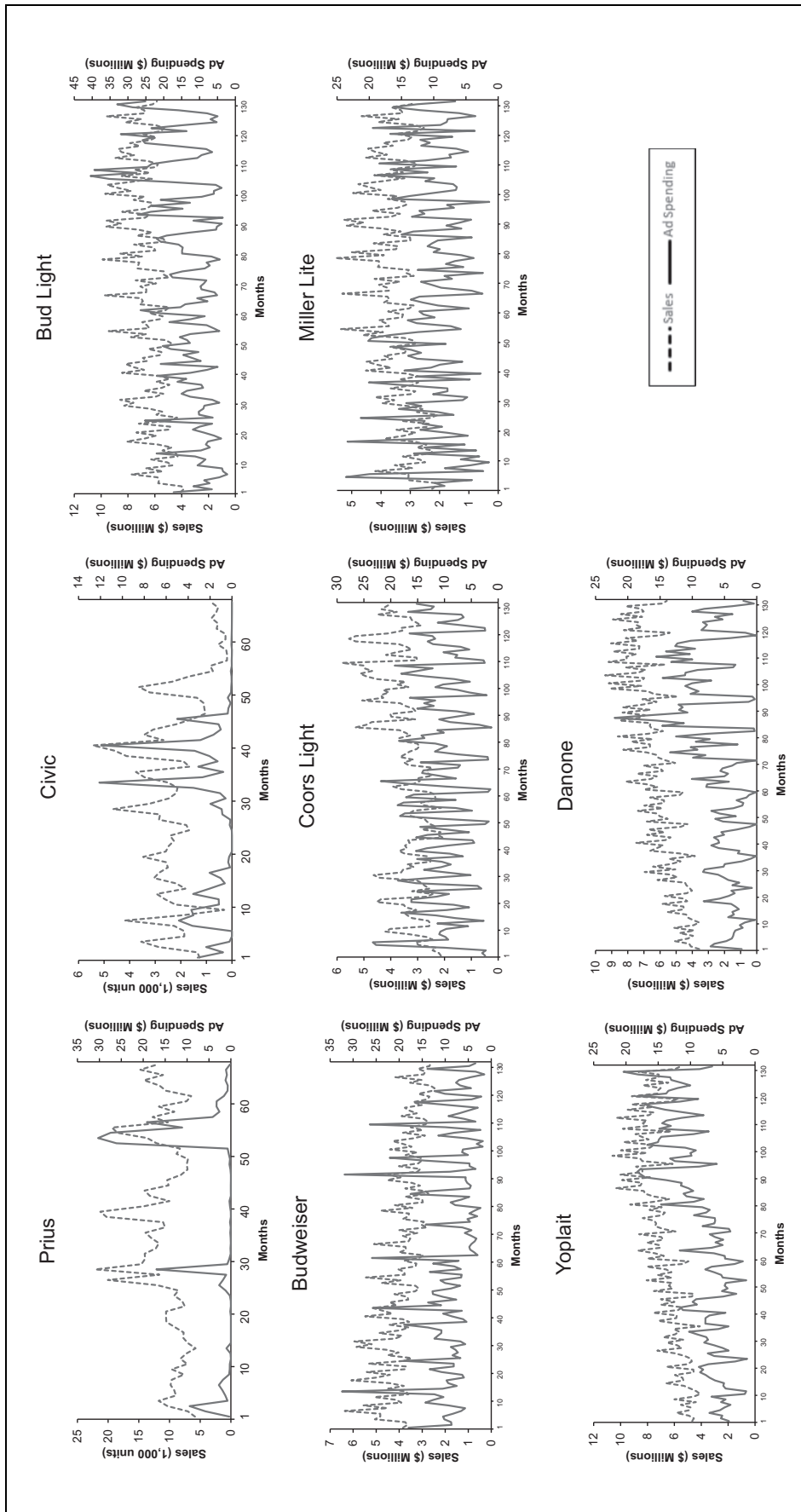


Figure 2. Sales and advertising over time.

**Parameter Estimates**

Table 5 reports parameter estimates and t-values (in parentheses) from the proposed model for the eight brands. Four key findings deserve attention. First, the majority of brands use a combination of methods in their advertising budgeting decisions, corroborating the practice paradigm of “multiple stakes in the ground.” However, they do so at varying degrees, as evidenced by the magnitudes of the corresponding parameter estimates, which are mostly significant at the 95% confidence level. Because the variables are scaled to similar operating ranges to achieve convergence in estimation, the magnitudes of the coefficients are comparable to some extent. Second, significant parameters have the expected signs. For example, promotions and coupons have positive effects on sales for the CPG brands. Similarly, price influences sales negatively by decreasing advertising spending and effectiveness. The direct effect of price on sales is insignificant, which is likely due to the limited variation (i.e., the median coefficient of variation is 3.23%) of the inflation-adjusted price variable akin to Liu and Shankar (2015). For hybrid cars, tax credits do not affect sales directly but boost the advertising effectiveness of Prius. Third, for many of the CPG brands, promotions and coupons have a negative impact on ad spending and tend to also have a negative impact on ad effectiveness. Thus, promotional activities not only reduce ad spending directly but also diminish ad effectiveness, which, in turn, reduces the optimal ad spending—a promotion “double whammy.”

Finally, Figure 3 displays time-varying advertising effectiveness and its corresponding 95% confidence intervals for all brands. The brands in the mature CPG categories (beer, yogurt) exhibit much less ad effectiveness uncertainty than the brands in the emerging niche of hybrid cars, as evidenced by the narrower confidence bands. Since CPGs and hybrid cars differ regarding their product life cycle stage, the resulting differences in ad effectiveness uncertainty are plausible. We use the proposed model and the Kalman filter estimation to quantify the time-varying uncertainty for each brand through the trajectory of  $\{\hat{\sigma}_{\beta_t}^2 : t = 1, \dots, N\}$ . Because ad effectiveness uncertainty is the premise for adaptive experimentation, we expect hybrid car brands to experiment more than the packaged goods brands.

**Model Validation**

To validate the proposed model, we first compare its performance with that of a benchmark model with no budget decomposition. Specifically, the benchmark model is specified by

$$\underbrace{\begin{bmatrix} S_t \\ \beta_{t+1} \end{bmatrix}}_{f_t} = \underbrace{\begin{bmatrix} \lambda & X_t \\ 0 & 1 \end{bmatrix}}_T \underbrace{\begin{bmatrix} S_{t-1} \\ \beta_t \end{bmatrix}}_{f_{t-1}} + \underbrace{\begin{bmatrix} d_{1t} \\ d_{2t} \end{bmatrix}}_{D_{1t}} + \underbrace{\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2,t+1} \end{bmatrix}}_{\varepsilon_t} \tag{12}$$

$$\underbrace{Y_t}_{y_t} = \underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_H \underbrace{\begin{bmatrix} S_t \\ \beta_{t+1} \end{bmatrix}}_{f_t} + v_t \tag{13}$$

In Equation 13, we delete the second row of Equation 7 to exclude budget decomposition, and we replace  $u_t$  in the first row of Equation 6 by  $X_t$  to obtain Equation 12. We set the training sample to 128 and 64 months for the CPG and hybrid categories, respectively, and we use the last four months as a holdout sample. Table 6 presents the results for both in- and out-of-sample performance. It reports the average one-step-ahead forecasts on the basis of the prior mean sales in the holdout window given the parameter estimates from the training sample. The proposed model shows an average 25.4% decrease in the mean absolute percentage prediction error (MAPE) compared to the benchmark model across all brands and an average 26.5% decrease in the root mean squared prediction error (RMSE). These decreases range from 1% to 56% for MAPE and from zero to 50% for RMSE. Also, the proposed model outperforms the benchmark model for every brand in all three categories. These findings provide strong evidence for the predictive validity of the proposed approach. Figure 4 displays the 95% confidence interval of the one-step-ahead predicted and actual sales. We observe that the actual sales generally lie within the confidence interval, garnering further support for the proposed model.

Finally, we assess multicollinearity using pairwise correlations between decision rules. Table 7 presents the median and range of pairwise correlations between budgeting methods for the eight brands. The average absolute correlation across the eight brands is .11. The two highest correlations are between competitive parity and the A/S heuristic for Yoplait (.61) and competitive parity and adaptive experimentation for Prius (−.65). However, as Table 5 shows, Yoplait’s estimated effect (.39) and t-value (2.41) for the A/S rule and the estimated effect (.52) and t-value (20.14) for the competitive parity rule are nonzero and significant. Similarly, for Prius, the estimated effect (1.40) and t-value (8.23) for adaptive experimentation and the estimated effect (.58) and t-value (119.44) for the competitive parity rule are nonzero and significant, thereby reassuring that collinearity does not substantively affect the empirical results.

**Implications**

**Ad Budget Decomposition**

To explore the insights from our findings, we convert the parameter estimates corresponding to the different decision rules in Table 4 into relative weights. In order to succinctly present the relative weights, we re-express the expected observed spending resulting from the second element of the observation vector in Equation 7 together with Equation 5 as follows:

$$y_{2t} = \hat{\gamma}_0 x_0 + \hat{\gamma}_1 x_{1t} + \hat{\gamma}_2 x_{2t} + \hat{\gamma}_3 x_{3t} + Z'_{3t} \hat{\delta}, \tag{14}$$

where  $y_{2t} = E[X_t]$  and  $x_0 = 1$  for all  $t$  so that the intercept  $\hat{\gamma}_0 = \hat{u}_b$  represents the baseline spending as before. Thus, the vector  $x_t = (1, \sigma_{\beta_t}^2, E_t, C_t)'$  represents the four analytic

**Table 5.** Estimation Results.

	Bud Light	Budweiser	Coors Light	Miller Lite	Yoplait	Danone	Prius	Civic
Carryover effect	.62*** (3.64)	.61** (2.36)	.54*** (2.93)	.64*** (1.87)	.53*** (1.73)	.53*** (1.88)	.94*** (27.64)	.41*** (1.91)
Initial ad effectiveness	.21 (1.1)	.93 (5.2)	.06 (0.2)	.17 (2.1)	.20 (0.9)	.21 (1.1)	.88 (3.9)	.45 (1.6)
Baseline spending	.21*** (108.33)	.03*** (70.47)	.06*** (29.27)	.01*** (76.48)	.02*** (19.04)	.02*** (15.04)	.10*** (5.51)	1.08 (.07)
Adaptive experimentation	.01*** (11.24)	.01*** (34.04)	.00*** (21.73)	.02*** (40.87)	.05*** (5.71)	.07*** (9.71)	1.40*** (8.23)	.37*** (9.96)
Advertising-to-sales heuristic	.08*** (5.56)	.23*** (21.25)	.15*** (5.93)	.08*** (7.83)	.39*** (2.41)	.33 (.46)	.02*** (39.69)	1.47*** (15.91)
Competitive heuristic	.09* (1.28)	.16*** (3.80)	.17** (1.85)	.40** (2.26)	.52*** (20.14)	.60*** (25.82)	.58*** (119.44)	.15** (1.48)
Price	-.27 (-4.5)	-.42 (-9.3)	-.28 (-1.4)	-.15 (-1.01)	-.21 (-2.6)	-.18 (-3.1)	-.05 (-.32)	-.07 (-.23)
Promotion	.55*** (3.74)	.10* (1.24)	-.00 (-0.2)	.03** (2.35)	.27*** (2.03)	.33*** (3.23)	.67 (.45)	1.31 (.02)
Coupon	.01*** (68.51)	.01*** (123.81)	.00*** (50.57)	.00*** (179.78)	.01*** (20.65)	.04*** (6.65)	.38 (.51)	.59 (.37)
Summer	.17*** (160.49)	.06*** (89.85)	.03*** (17.47)	.04*** (78.69)	-.10*** (-8.94)	-.03*** (-2.17)	1.70* (1.25)	1.42* (1.56)
Spring	.13*** (161.57)	.11*** (289.88)	.02*** (14.49)	.02*** (82.84)	.02*** (5.03)	-.01 (-.59)	.05 (0.04)	.09 (.14)
Fall	-.04*** (-26.16)	-.03*** (-32.45)	-.02*** (-45.85)	-.00*** (-26.57)	-.10*** (-16.05)	-.00 (-.03)	-.08*** (-4.19)	-.05 (-1.06)
Price	-.67* (-1.47)	-2.09*** (-2.56)	-.38 (-2.1)	-1.56* (-1.31)	-30*** (-69.71)	-2.9** (-2.12)	.33*** (2.11)	.33 (.23)
Promotion	-.29*** (-7.48)	.01 (.39)	-.05*** (-3.07)	.19*** (6.17)	-.11*** (-132.69)	-.12*** (-7.26)	.08 (.75)	.30** (2.06)
Coupon	-.02*** (-72.97)	-.00*** (-11.08)	-.01*** (-72.52)	.00*** (49.62)	-.00*** (-124.05)	-.01*** (-19.66)	.13 (.72)	-.13 (-.55)
Summer	-.11*** (-3.18)	-.06** (-1.24)	-.12*** (-3.16)	-.07** (-1.28)	.07*** (47.55)	.04*** (4.85)	.24** (2.29)	-.12 (-.17)
Spring	.50*** (22.76)	.70*** (16.89)	.36*** (11.91)	.37*** (53.29)	.04*** (43.35)	.04*** (7.88)	-1.70*** (-14.15)	-2.74 (-.38)
Fall	.21*** (9.97)	.24*** (7.54)	-.01 (-.39)	.17*** (11.27)	-.02*** (-6.31)	.02*** (1.71)	-.66** (-1.67)	.95 (.04)
Price	-.02 (-0.1)	.47 (.30)	.01 (.00)	-.31 (-.15)	-.02 (-.03)	-.05 (-.00)	.58*** (3.61)	.45 (.50)
Promotion	-.74*** (-41.80)	-.18*** (-7.57)	-.05** (-1.91)	2.1*** (5.84)	.16** (1.53)	.17 (.64)	-.43 (-.86)	42 (.32)
Coupon	-.01*** (-102.48)	-.00*** (-188.94)	-.01 (-.05)	-.00*** (254.63)	-.04*** (-71.63)	-.02*** (-3.50)	-.33* (-1.14)	.59 (.40)
Summer	-.10*** (-31.28)	-.05*** (-349.87)	-.01*** (-20.55)	-.02*** (-251.32)	.09*** (21.18)	.09*** (21.55)	.21 (.01)	5.44*** (6.55)
Spring	-.05*** (-166.62)	-.06*** (-405.95)	.01*** (36.64)	.03*** (282.50)	-.09*** (-31.65)	.21*** (48.60)	.41*** (163.49)	.46*** (78.27)
Fall	-.02*** (-61.84)	-.04*** (-251.41)	.03*** (68.56)	.02*** (150.93)	.08*** (26.65)	-.21*** (-49.99)	3.15*** (31.45)	1.54*** (40.02)
Process noise	5e-5** (1.87)	7e-5*** (7.74)	4e-5*** (2.23)	4e-4 (.00)	1e-5 (.00)	5e-5*** (1.96)	-.238.98	-.125.02
Measurement noise	.07*** (655.53)	.05*** (106.50)	.05*** (104.28)	.04*** (12.72)	.20*** (16.80)	.23*** (172.50)		
Log likelihood	557.51	662.07	672.83	690.07	296.38	249.93		

Notes: t-values in parentheses. \*p = .10, \*\*p = .05, \*\*\*p = .01. Annual controls are omitted from the table for brevity.

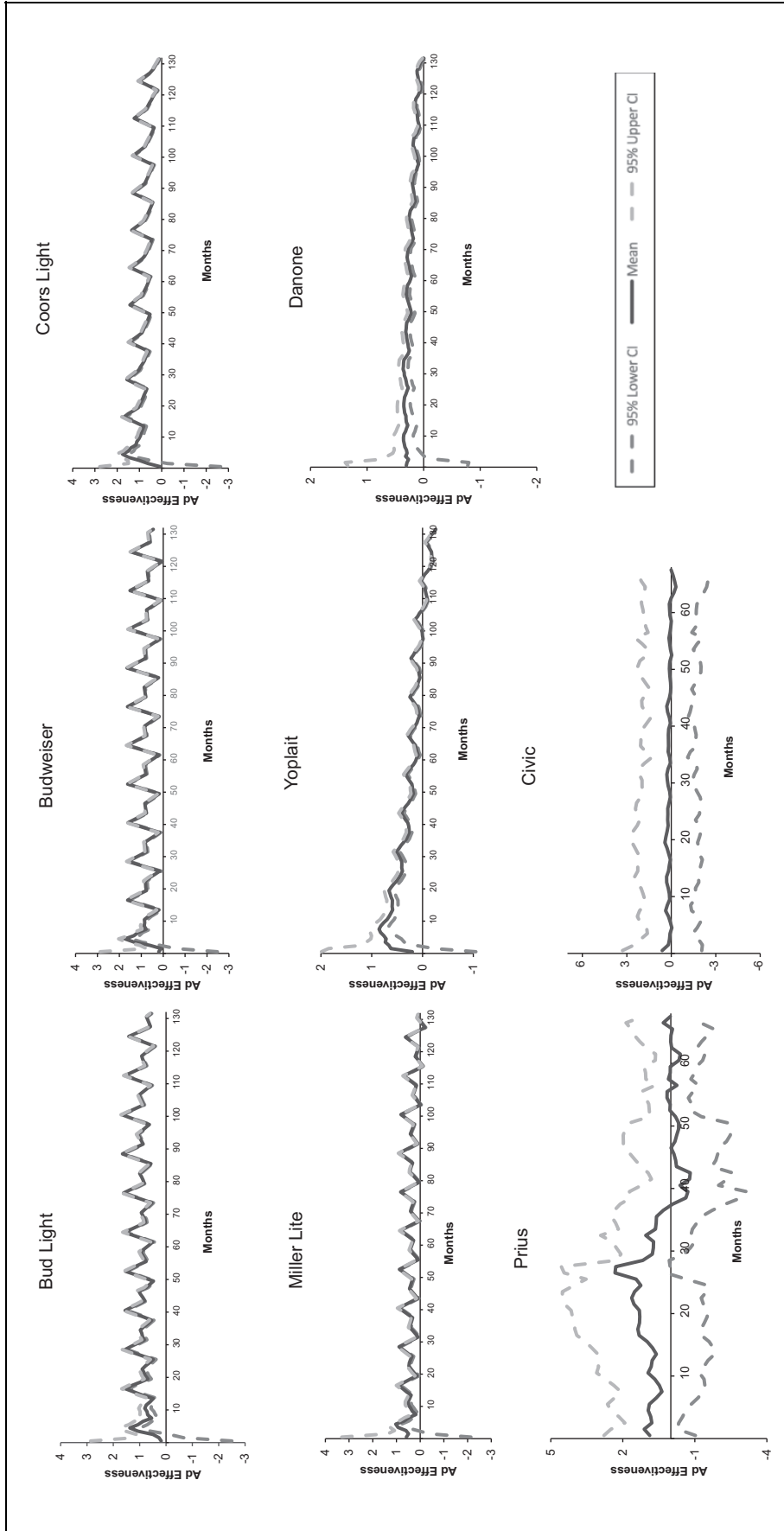


Figure 3. Time varying ad effectiveness and its uncertainty.

**Table 6.** Model Validation.

Brand	In-Sample Performance			Holdout Performance		
	Performance Metric	Benchmark Model	Proposed Model	Performance Metric	Benchmark Model	Proposed Model
Bud Light	AIC	-1032.70	-1057.80	MAPE (%)	14.70	8.05
	BIC	-960.63	-973.42	RMSE	.10	.05
Budweiser	AIC	-1250.35	-1266.14	MAPE (%)	7.19	4.90
	BIC	-1178.28	-1182.54	RMSE	.02	.01
Coors Light	AIC	-1221.60	-1287.66	MAPE (%)	16.47	7.10
	BIC	-1149.23	-1204.06	RMSE	.06	.03
Miller Lite	AIC	-1302.00	-1322.00	MAPE (%)	7.93	5.29
	BIC	-1229.93	-1238.40	RMSE	.03	.02
Yoplait	AIC	-504.60	-534.76	MAPE (%)	14.61	12.86
	BIC	-432.23	-451.16	RMSE	.10	.10
Danone	AIC	-419.56	-441.86	MAPE (%)	11.12	9.77
	BIC	-347.49	-358.26	RMSE	.07	.07
Prius	AIC	612.36	523.96	MAPE (%)	16.75	14.98
	BIC	665.91	590.26	RMSE	2.27	1.92
Civic	AIC	311.84	296.04	MAPE (%)	55.92	55.49
	BIC	366.61	362.34	RMSE	.40	.35

Notes: Holdout results report one-step-ahead forecasts for the last four months in the observation window. MAPE denotes the mean absolute percentage error over the forecast period. AIC and BIC denote Akaike and Bayesian information criteria, respectively.

and heuristic methods. We then calculate the relative weight for each method as follows:

$$w_i = \frac{\hat{\gamma}_i \bar{x}_i}{\sum_{i=0}^3 \hat{\gamma}_i \bar{x}_i}, \quad (15)$$

where  $\bar{x}_i$  denotes the sample averages for  $i = 0, 1, 2, 3$  and  $\gamma_i \bar{x}_i$  represents the level importance of the corresponding method (Achen 1982, p. 72), which Kruskal and Majors (1989) found to be a popular choice in their survey of scientific publications. More specifically, they note, “An approach popular with economists for inherently positive variates is to take relative importance proportional to  $\gamma_i \bar{x}_i, \dots$ ” (p. 5; notation adapted to our context). The reason for the popularity with economists is that the level importance is proportional to elasticity. Indeed, for  $i = 1, 2, 3$ ,  $\hat{\gamma}_i \bar{x}_i \propto \eta_i = \frac{\bar{x}_i}{\bar{y}_2} \frac{\partial \bar{y}_2}{\partial x_i} = \frac{\hat{\gamma}_i \bar{x}_i}{\bar{y}_2}$ , where  $\eta_i$  is the spending elasticity with respect to  $x_i$  evaluated at the mean levels (and  $\bar{y}_2$  is constant over  $i$ ). An added benefit of using level importance is that the weights incorporate the relative contribution of the intercept (Achen 1982), which in our case represents the baseline spending ( $i = 0$ ). Thus, consistent with the discussion in Achen (1982) and Kruskal and Majors (1989), the relative weights in Equation 15 depend on the level importance  $\hat{\gamma}_i \bar{x}_i$ . Table 8 reports the relative contributions of each method to the ad budgeting decision, calculated according to Equation 15 and expressed in percentages.

The results shed light on how managers combine the different analytics and heuristics-based methods to determine their ad spending and how such decisions vary across brands and categories. For each brand, we also report advertising elasticities, which range from .07 to .27 and broadly comport with those reported in prior literature (Assmus, Farley, and Lehmann 1984; Sethuraman, Tellis, and Briesch 2011).

Table 8 indicates that the combinations of budgeting rules vary across brands and categories. Three key observations emerge: First, car brands tend to use experimentation to a greater extent because hybrid cars represent an emerging niche characterized by an uncertain market environment with limited knowledge of consumer response to a new technology and/or brand advertising. Second, yogurt brands use almost exclusively heuristics-based methods, maintain a similar balance between A/S and competitive benchmarking, and exhibit similar advertising elasticities. Third, beer brands use a mix of approaches, with the leading brand (Bud Light) relying on baseline spending, Coors Lite using a balance of analytics and heuristics methods, and the other two (Budweiser and Miller Lite) relying on heuristics but with different combinations of the A/S and competitive benchmarking methods.

Most importantly, ad budget decomposition not only offers new insights for firms’ internal audits of their own brands and business intelligence on competing brands but it also reveals information at the brand level on the basis of market data, which would be impossible using the survey-based approaches in the extant literature (e.g., Farris and West 2007) that summarize at the *industry* level the frequency counts of managers’ opinions on which methods they use in their companies. More specifically, our findings suggest that budgeting decisions are brand-specific, and one size does not fit all as far as the budgeting mix is concerned. Rather, different methods and combinations are deployed by different brands, possibly to leverage conditions such as market leadership, market maturity, and category growth. For example, car brands competing in an emerging niche tend to use more experimentation than mature CPG brands due to the higher uncertainty characterizing advertising effectiveness. In terms of heuristics, Prius relies considerably on competitive parity to maintain its leadership, whereas

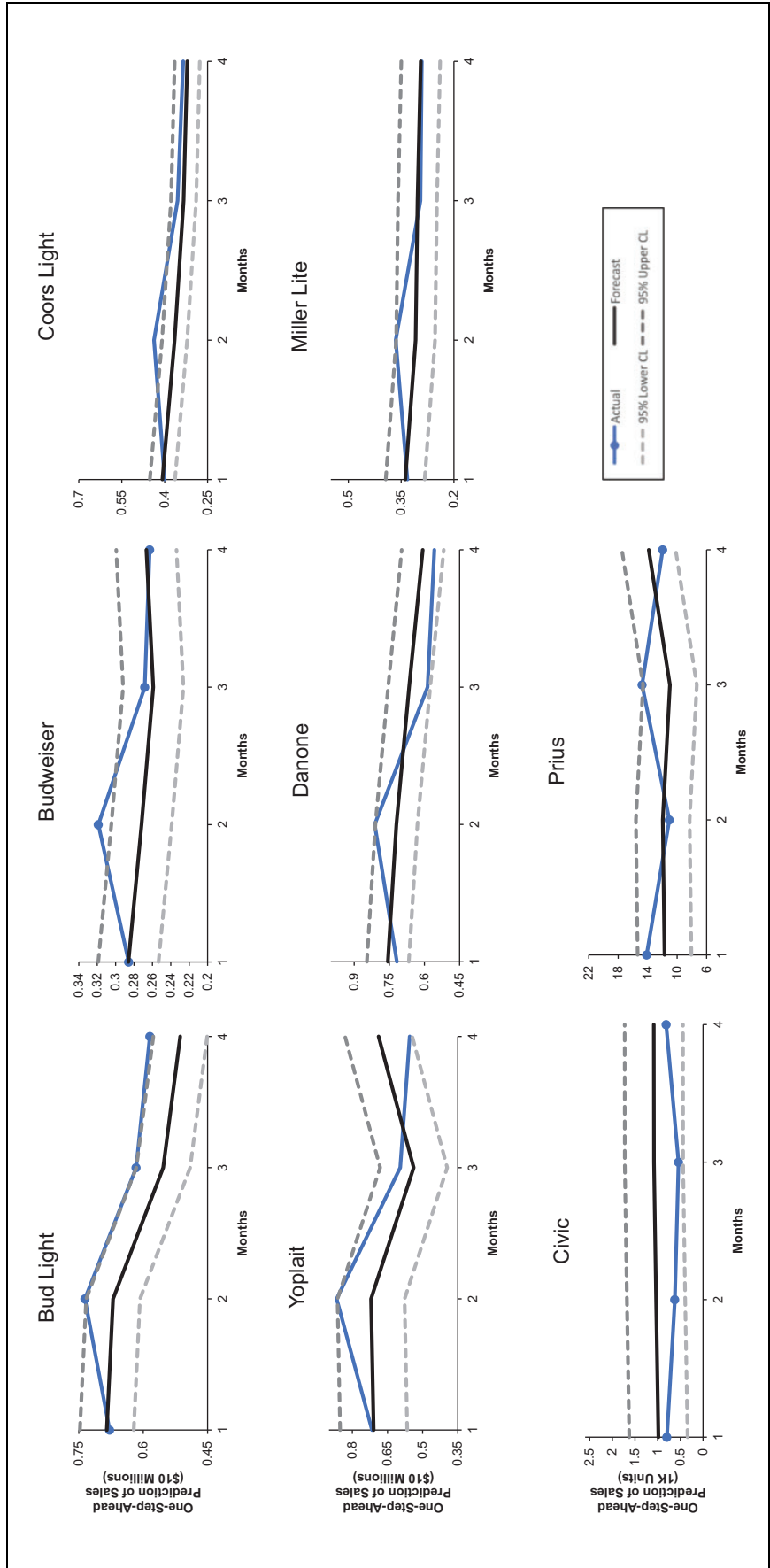


Figure 4. One-step-ahead forecasts.

**Table 7.** Correlations Between Budget Decision Rules.

	Baseline Spending		Experimentation		Advertising-to-sales Heuristic	
	Median	Range	Median	Range	Median	Range
Experimentation	.00	(.00, .00)				
Advertising-to-sales heuristic	.00	(.00, .00)	-.19	(-.37, .13)		
Competitive parity heuristic	.00	(.00, .00)	-.07	(-.65, .13)	.23	(-.02, .61)

Notes: Figures across eight brands.

**Table 8.** Relative Contributions.

		Relative Contribution to the Ad Budget Decision							
		Bud Light	Budweiser	Coors Light	Miller Lite	Yoplait	Danone	Prius	Civic
Analytics-based methods	Baseline spending	78%	16%	52%	17%	4%	11%	1%	19%
	Adaptive experimentation	1%	1%	1%	1%	1%	1%	20%	41%
	Total analytics	79%	17%	53%	18%	5%	12%	21%	60%
Heuristic-based methods	A/S ratio	19%	78%	43%	56%	53%	34%	33%	33%
	Competitive parity	2%	5%	4%	26%	42%	54%	46%	7%
	Total heuristics	21%	83%	47%	82%	95%	88%	79%	40%
	Advertising elasticity	.18	.16	.27	.20	.21	.22	.08	.07

Notes: Percentages rounded to nearest integers.

Civic relies on the A/S ratio method, possibly to leverage growth and strengthen its market position (see also Wang and Zhang 2008). Similarly, the yogurt brands, which enjoyed growth over the observation period, rely on A/S to leverage market growth conditions.

### Evidence for Adaptive Experimentation

A primary goal of this study was to empirically test the hitherto untested hypothesis—that experimentation budget is proportional to the uncertainty in ad effectiveness—grounded in the dual control theory (see Filatov and Unbehauen 2004). This study is the first to furnish strong empirical evidence. Theoretically, we expected ad budgets to increase proportionately to the uncertainty in ad effectiveness, or  $\partial u_t / \partial \sigma_{\beta_t}^2 = \gamma_1$  to be positive. Managerially, this expectation means that a larger budget for adaptive experimentation should be set aside when managers expect greater uncertainty in ad effectiveness (e.g., when launching new ad campaigns or products). Empirically, we find that  $\hat{\gamma}_1$  is uniformly positive and statistically significant for all eight brands from three different categories (see Table 5). The results demonstrate that managers spend ad dollars proportionately to the uncertainty in ad effectiveness in durable and nondurable categories, contributing pioneering evidence to support adaptive experimentation in marketing. Consistent with our expectations, allocation due to experimentation is higher in the emerging hybrid car niche due to the higher advertising effectiveness uncertainty.

### Revenue Implications of Shifting to Adaptive Experimentation

Would more experimentation benefit brands? To answer this question, we conduct simulations to reallocate heuristic

spending to experimentation. Specifically, we shift 50% of the ad budget due to competitive benchmarking to experimentation for each brand and calculate the change in average sales over the observation period. In the process, we kept the ad budget for each year fixed to its observed value and performed the budget reallocation on an annual basis. Here we present the key steps of this simulation:

1. Calculate the estimated contribution of the four decision rules to the advertising budget:  $\hat{u}_b$  for baseline spending,  $\hat{\gamma}_1 \hat{\sigma}_{\beta_t}^2$  for experimentation,  $\hat{\gamma}_2 E_t$  for A/S ratio, and  $\hat{\gamma}_3 C_t$  for competitive parity;
2. Calculate the aggregate contribution of the competitive parity heuristic for each year and reallocate 50% of it to experimentation;
3. For each year, assign the annual shift calculated in Step 2 to experimentation for each month weighted by the experimentation spending of that month (i.e., each month's additional allocation is proportional to that month's experimentation contribution to the annual experimentation contribution);
4. Evaluate the new total predicted sales from the simulated contributions and calculate the percent change from the total predicted sales prior to the reallocation.

Table 9 presents the simulation results. All brands experience a lift in sales due to the increased experimentation spending. The change is modest for brands in the mature categories, ranging from .97% to 1.89%. In contrast, the hybrid car brands benefit substantially, as evidenced by the double-digit sales lifts. The incremental benefit of experimentation is lower for Civic, at 12%, than the market leader Prius, at 23%. However, the shift



**Table 9.** Shifting Allocation from Heuristics to Adaptive Experimentation.

	Bud Light	Budweiser	Coors Light	Miller Lite	Yoplait	Danone	Prius	Civic
% change in sales performance	.97	1.04	1.89	.99	1.32	1.80	23.24	11.89
Shift as % of the total ad budget	1.09	2.49	2.21	13.27	23.80	27.10	33.43	14.44
Average monthly shift (\$ million)	.15	.22	.31	1.32	2.07	1.66	.92	.20

represents 14% of the total budget for Civic versus 33% for Prius, thereby explaining the differential returns to experimentation between the two brands. These findings not only corroborate the arguments of dual control theory regarding the role of uncertainty in ad effectiveness but also show there is room for improvement by relying on more analytics to determine advertising budgets.

## Conclusion

Our analysis reveals that advertising budgeting is a complex process requiring managers to make decisions under uncertainty of ad effectiveness. To assess ad effectiveness, managers may undertake field experiments by manipulating ad spending over a fixed duration in chosen test/control regions. However, advertising field experiments are not common in practice because they typically lack the power to accurately measure ad effectiveness. Informative experiments can require more than 10 million person-weeks in test/control markets, making them costly and infeasible for many firms (Lewis and Rao 2015, p. 1941). More specifically, on the basis of 25 large field experiments, Lewis and Rao (2015, p. 1950) note that "... even a very successful campaign has a minuscule  $R^2$  of 0.0000054," and "... reliably distinguishing a 50% from 0% ROI is typically not possible with a \$100,000 experiment involving millions of individuals" (ibid., p. 1964). Even when undertaken, experiments eventually come to an end and business-as-usual sets in, leaving managers guessing the time-varying magnitudes of ad effectiveness and the level of uncertainty associated with it. Thus, we first fill this gap in the literature to provide managers with a framework to learn how both ad effectiveness and its uncertainty varies over time for their own sales-advertising data.

Second, according to our analyses of eight brands from three categories, durable and nondurable at different stages of the product life cycle, managers can learn how to reverse engineer their budgeting processes by decomposing ad budgets into four methods: baseline spending, adaptive experimentation, advertising-to-sales heuristic, and competitive parity heuristic. This decomposition quantifies the relative contributions of each method, corroborating the "multiple stakes in the ground" paradigm of ad budgeting, which has so far been captured through surveys aggregating information at the industry level and relying on managers' opinions (e.g., Farris and West 2007). In contrast, using our proposed model, managers can learn about the decomposition (1) of both their own (i.e., self-audit) and competing brand ad budgets (i.e., business intelligence) in their industry, (2) at the brand level rather than the aggregated

industry level, and (3) based on market data, not managers' opinions.

Third, we find that budgeting decisions are brand-specific even for brands belonging to the same corporate group (e.g., beer brands), which is consistent with previous empirical evidence provided by Anheuser-Busch experiments (Ackoff and Emshoff 1975) and Kraft's product harm crisis in Australia (Van Heerde, Helsen, and Dekimpe 2007). Thus, our findings reflect the consideration of brand-specific needs in budget setting for multi-brand companies (Low and Mohr 1999) and the role of brand architecture (see Avery 2016, pp. 12–17; Hill, Ettenson, and Dane Tyson 2005). More specifically, due to their different market positioning and objectives, brands belonging to the same architecture of a multi-brand firm require different budgeting approaches, which may also be dictated by their market performance. We observe this reality in the beer category, in which Bud Light and Budweiser, the top selling brands of Anheuser-Busch, use different budgeting methods. Bud Light, the leading brand in a category in which ad effectiveness uncertainty is low, heavily relies on baseline spending to maintain its market leadership (Wang and Zhang 2008). In contrast, Budweiser, the number two brand in the Anheuser-Busch portfolio, experienced declining sales during the observation period, and its budget is determined predominantly by the A/S ratio heuristic. This could reflect an effort to make advertising spending commensurate with sales performance to justify advertising support.

Finally, pioneering articles of Little (1966) and Pekelman and Tse (1980) advocated the concept of adaptive experimentation, which was later grounded in the dual control theory (see Filatov and Unbehauen 2004), suggesting that the experimentation budget should be proportional to the uncertainty in ad effectiveness. This hypothesis was hitherto untested in the marketing literature. This study is the first one to provide empirical evidence that shows managers spend ad dollars proportional to the uncertainty in ad effectiveness, shedding light on the rationality of managers departing from profit-maximizing optimality. In other words, suboptimal behavior does not imply nonrational behavior.

Much more can be discovered by extending this study. Replication across many brands and multiple industries would reveal how relative weights depend on market conditions (e.g., scope) and brand positioning (e.g., role). Extension of the total budget split across analytics and heuristics subcomponents to each of the multimedia channels may yield novel insights. Finally, an empirical demonstration that a combination of different budgeting rules is the outcome of joint negotiations between multiple actors in the decision-making unit

would make an important contribution to the canons of advertising.

### Appendix: Stochastic Optimal Control of Advertising

#### State Dynamics

Equation 6 specifies the evolving sales and time-varying parameters in the state vector  $f_t = (S_t, \beta_{t+1})'$

$$\begin{bmatrix} S_t \\ \beta_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda & u_t \\ 0 & 1 \end{bmatrix}}_{T_t} \underbrace{\begin{bmatrix} S_{t-1} \\ \beta_t \end{bmatrix}}_{f_{t-1}} + \underbrace{\begin{bmatrix} d_{1t} \\ d_{2t} \end{bmatrix}}_{D_{1t}} + \underbrace{\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2,t+1} \end{bmatrix}}_{\varepsilon_t}$$

where  $T_t$  denotes the transition matrix, the drift vector  $D_{1t} = (d_{1t}, d_{2t})'$ , the error terms  $(\varepsilon_{1t}, \varepsilon_{2,t+1})' = \varepsilon_t \sim N(0, Q)$ , and  $Q = \sigma_\varepsilon^2 I_2 \times 2$ .

First, we apply algebra to transform the discrete time equations to their continuous time versions to facilitate the derivation of the optimal policy. From the sales dynamics  $S_t = \beta_t u_t + \lambda S_{t-1} + d_{1t} + \varepsilon_{1t}$ , we subtract  $S_{t-1}$  from both the sides and note that  $\Delta t = 1$  to obtain  $\frac{\Delta S}{\Delta t} = S_t - S_{t-1} = \beta_t u_t - (1 - \lambda) S_{t-1} + d_{1t} + \varepsilon_{1t}$ . Thus,  $dS = \frac{\Delta S}{\Delta t} dt = [\beta_t u_t - \delta S_{t-1} + d_{1t}] dt + \varepsilon_{1t} dt$ , where  $\delta = (1 - \lambda)$ .

Next, how do we represent “shocks” in continuous time? Following Sethi and Thompson (2000, pp. 343–345), we proceed in two steps. First, we invoke the standard “white noise process,” denoted by  $n_1(t)$ , which is Gaussian at each  $t$ , independent at  $t \neq t'$  (i.e., zero autocorrelation), and use it to replace the independent normal random sequence  $\varepsilon_{1t}$ . The term “sequence” refers to the discrete time series, and “process” refers to its continuous time version, by convention. Then we re-express  $dS = [\beta_t u_t - \delta S_{t-1} + d_{1t}] dt + \sigma_\varepsilon n_1(t) dt$ . The standard deviation  $\sigma_\varepsilon$  multiplies  $n_1(t)$  because  $n_1(t)$  has a unit variance (i.e., standard white noise akin to the standard normal), a  $\varepsilon_{1t}$  follows  $N(0, \sigma_\varepsilon^2)$ . In the second step, we convert the white noise process to the Wiener process to get a stochastic differential equation in the Ito sense. Specifically, we replace the white noise process  $n_1(t)$  with the Brownian motion  $B_1(t)$ , also known as the Wiener process, to obtain the Ito stochastic differential equation for sales:

$$dS = (\beta u - \delta S + d_1) dt + \sigma_\varepsilon dB_1. \tag{A1}$$

In Equation A1 and elsewhere in the Appendix, the variables  $(S, \beta, u, d_1, d_2, B_1, B_2)$  are functions of time  $t$ . By comparing  $n_1(t) dt$  and  $dB_1$ , we can think of the white noise process  $n_1(t)$  as the generalized derivative of the Wiener process  $dB_1/dt$ , which formally is differentiable nowhere but continuous everywhere. Likewise, we derive the continuous time stochastic differential equation for ad effectiveness:

$$d\beta = d_2 dt + \sigma_\varepsilon dB_2, \tag{A2}$$

where  $(\beta, d_2, B_2)$  are functions of time  $t$ . We suppress  $t$  to improve readability and emphasize it when needed for clarity.

Equations A1 and A2 capture both the dynamics and the uncertainty at each instant  $t$  in sales and ad effectiveness, respectively.

#### Net Present Value

When we view  $u(t)$  as a trajectory over the span  $t \in [0, \infty)$ , what is the net present value associated with *any* ad spending trajectory  $u(t)$ ? To answer this question, we compute the expected total discounted profit as

$$\underbrace{J(u)}_{\text{NPV}} = \mathbb{E} \left[ \underbrace{\int_0^\infty \underbrace{e^{-\rho t}}_{\text{discount factor}} \underbrace{(mS - c(u)) dt}_{\text{instantaneous profit}}}_{\text{discounted profit}} \right] \tag{A3}$$

cumulative profit

Expected total discounted profit

where  $J(u)$  denotes the net present value,  $u(t)$  is ad spending trajectory over time,  $\rho$  represents the discount rate, and  $m$  equals the contribution margin (i.e., price minus variable costs). The term  $mS - c(u)$  represents the earnings at instant  $t$ . To this end, sales should be measured in dollar revenues when using a percentage margin and in unit sales when using a contribution margin. The discount factor  $e^{-\rho t}$  captures the time value of money: earnings at  $t$  is more than the same earnings at a future time  $t' > t$ . The integral accumulates the discounted earnings over time, and the expectation averages over the ensemble of stochastic trajectories due to Equations A1 and A2. Thus,  $J(u)$  is the expected net present value (NPV) of any of the infinitely many advertising plans  $u(t)$ .

#### Optimal Advertising Strategy

Now consider a particular spending trajectory  $u_1(t)$  that results in the specific NPV  $J_1 = J(u_1(t))$ . Consider an alternative trajectory, say  $u_2(t)$ , that generates NPV  $J_2 = J(u_2(t))$ . Let various advertising trajectories  $\{u_1(t), u_2(t), \dots, u_\infty(t)\}$  denote the set of the infinite number of feasible trajectories. Given the resulting net present values  $\{J_1, J_2, \dots, J_\infty\}$ , the optimal trajectory  $u^*(t)$  is the one that corresponds to the largest NPV  $J^* = J(u^*(t))$ .

This problem of finding  $u^*(t)$  is known as the stochastic control problem. It has remained unsolved in the marketing literature as yet because the state dynamics in Equations A1 and A2 are stochastic due to the Wiener processes (recall that Raman et al. [2012] ignore uncertainty). Thus, we seek to maximize the NPV  $J(u)$  in Equation A3 with respect to  $u(t)$  subject to the stochastic dynamic constraints in Equations A1 and A2.

To ensure global maximum, we employ a quadratic  $c(u) = u^2$ , which has been used in the advertising literature over the past 40 years (e.g., Chintagunta and Jain 1992; Little 1977; Naik, Prasad, and Sethi 2008; Rubel and Prasad 2016).

Let the optimal advertising trajectory  $u^*(t)$  attain the maximum NPV, which we refer to as the value function  $V(S, \beta)$ . Formally, the maximum NPV is given by the value function  $V(S, \beta) = \underset{u(t)}{\text{Max}} J(u(t))$ . The value function depends on the

state of the stochastic dynamic system, namely, sales  $S$  and ad effectiveness  $\beta$ . According to the stochastic control theory (e.g., Chapter 13 in Sethi and Thompson 2000), the value function  $V$  satisfies the Hamilton-Jacobi-Bellman equation:

$$\rho V = \underset{u}{\text{Max}} \left\{ (mS - u^2) + V_S(\beta u - \delta S + d_1) + V_\beta(d_2) + \frac{1}{2} V_{SS} \sigma_\epsilon^2 + \frac{1}{2} V_{\beta\beta} \sigma_\epsilon^2 \right\} \quad (\text{A4})$$

where  $V_S = \partial V / \partial S$ ,  $V_\beta = \partial V / \partial \beta$ ,  $V_{SS} = \partial^2 V / \partial S^2$ , and  $V_{\beta\beta} = \partial^2 V / \partial \beta^2$ .

The second partial derivatives  $V_{\beta\beta}$  and  $V_{SS}$  make Equation A4 a second-order differential equation. To solve this second-order differential equation, we apply Lagrange's method of undetermined coefficients. Generalizing Raman et al. (2012) to two state variables, we conjecture  $V(S, \beta) = a + bS + p\beta + q\beta^2$ . By differentiating the terms in the parenthesis of Equation A4 with respect to  $u$  and setting the result to zero, we obtain the first-order condition:

$$-2u + V_S \beta \equiv 0 \Rightarrow u^* = V_S \beta / 2. \quad (\text{A5})$$

Substituting Equation A5 in Equation A4 and replacing  $V_S = b$ ,  $V_\beta = p + 2q\beta$ ,  $V_{SS} = 0$ , and  $V_{\beta\beta} = 2q$ , we obtain:

$$\begin{aligned} & \rho(a + bS + p\beta + q\beta^2) \\ &= mS - \left(\frac{b\beta}{2}\right)^2 + b\left(\beta\left(\frac{b\beta}{2}\right) - \delta S + d_1\right) \\ & \quad + (p + 2q\beta)d_2 + q\sigma_\epsilon^2 \\ &= (bd_1 + pd_2 + q\sigma_\epsilon^2) + (m - b\delta)S + (2qd_2)\beta \\ & \quad + \left(\frac{b^2}{4}\right)\beta^2. \end{aligned}$$

By equating the corresponding terms on both sides, we explicitly solve for  $(a, b, p, q)$  in terms of model parameters. For instance,  $\rho bS = (m - b\delta)S$  yields  $b = \frac{m}{\rho + \delta}$ . This completes the derivation, leading to the following:

**P:** The optimal advertising in the presence of dynamic uncertainty in sales and ad effectiveness is given by  $u_t^* = \frac{m\beta_t}{2(\rho + \delta)}$ .

Having derived the optimal advertising in closed form, we can obtain the comparative statics with respect to parameters, which we omit for brevity but illustrate through an example to show how doing so provides insight: The optimal advertising  $u_t^*$  increases as the carryover effect  $\lambda$  increases. To substantiate,  $\frac{\partial u_t^*}{\partial \lambda} = \frac{\beta_t}{2} \frac{\partial \mu}{\partial \lambda} = \frac{m\beta_t}{2(\rho + \delta)^2} > 0$  because  $\delta = 1 - \lambda$ . The managerial intuition behind this forward-looking insight is as follows: The sales impact of current advertising persists in the

future due to the carryover effect, which in turn generates additional sales in the future that, in turn, justify increasing advertising spends today.

We close by noting that, to use the proposition in practice, managers should set up an accounting system consistent with the generally accepted accounting principles, which entails two parts: (1) computing the short-term contribution margin at the brand level such that the variable cost includes both the direct and allocated cost of goods sold across multiple brands, assigning the costs of shared components (e.g., common inventory carrying and transportation costs, shared costs of fulfillment centers, shared raw materials such as electricity, water, effluent treatment, common parts, property, plant, and equipment) to each of the various brands in the company's portfolio and (2) computing the weighted average cost of capital  $\rho$  by assigning the various costs of capital (e.g., debt coverage, dividend payout) to each of the various brands. In the absence of such an accounting system, the brand's long-term margin and thus the optimal advertising budget would remain unknown.

## Associate Editor

Peter Danaher


## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iD

Ceren Kolsarici  <https://orcid.org/0000-0003-4784-6903>

## References

- Achen, Christopher H. (1982). *Interpreting and Using Regression*. Beverly Hills, CA: SAGE Publications.
- Ackoff, Russell L. and James R. Emshoff (1975), "Advertising Research at Anheuser-Busch, Inc. (1963-68)," *Sloan Management Review (Winter)*, 16 (2), 1-15.
- Assmus, Gert, John U. Farley, and Donald R. Lehmann (1984), "How Advertising Affects Sales: Meta-Analysis of Econometric Results," *Journal of Marketing Research*, 21 (1), 65-74.
- Ataman, M. Berk, Harald J. van Heerde, and Carl F. Mela (2010), "The Long-Term Effect of Marketing Strategy on Brand Sales," *Journal of Marketing Research*, 47 (5), 866-82.
- Avery, Jill (2016), "Brand Portfolio Strategy and Brand Architecture," Case No. 517021. Boston, MA: Harvard Business School.
- Baumol, William J. (1962), "On the Theory of Expansion of the Firm," *The American Economic Review*, 52 (5), 1078-87.
- Bigné, J. Enrique (1995), "Advertising Budget Practices: A Review," *Journal of Current Issues & Research in Advertising*, 17 (2), 17-31.

- Bronnenberg, Bart J., Michael W. Kruger, and Carl F. Mela (2008), "Database Paper—The IRI Marketing Data Set," *Marketing Science*, 27 (4), 745–48.
- Bruce, Norris (2008), "Pooling and Dynamic Forgetting Effects in Multi-Theme Advertising: Tracking the Ad Sales Relationship With Particle Filters," *Marketing Science*, 27 (4), 659–73.
- Chintagunta, Pradeep K. and Dipak Jain (1992), "A Dynamic Model of Channel Member Strategies for Marketing Expenditures," *Marketing Science* 11 (2), 168–88.
- Danaher, Peter J., André Bonfrer, and Sanjay Dhar (2008), "The Effect of Competitive Advertising Interference on Sales for Packaged Goods," *Journal of Marketing Research*, 45 (2), 211–25.
- Danaher, Peter J. and Ronald Rust (1994), "Determining the Optimal Level of Media Spending," *Journal of Advertising Research*, 34 (1), 28–35.
- Danenberg, Nick, Rachel Kennedy, Virginia Beal, and Byron Sharp (2016), "Advertising Budgeting: A Reinvestigation of the Evidence on Brand Size and Spend," *Journal of Advertising*, 45 (1), 139–46.
- Davidson, Russell and James G. MacKinnon (2004), *Econometric Theory and Methods*. New York: Oxford University Press.
- Eastlack, Joseph O. and Ambar G. Rao (1989), "Advertising Experiments at the Campbell Soup Company," *Marketing Science*, 8 (1), 57–71.
- Farris, Paul W. and Douglas C. West (2007), *A Fresh View of the Advertising Budget Process*. London: SAGE Publications.
- Feldbaum, Aleksandr A. (1965), *Optimal Control Systems*. New York: New York Academic Press.
- Filatov, Nikolai M. and Heinz Unbehauen (2004), *Adaptive Dual Control: Theory and Applications*. Verlag Berlin Heidelberg, Germany: Springer.
- Gunny, Katherine (2005), "What Are the Consequences of Real Earnings Management?" ProQuest Dissertations Publishing, University of California, Berkeley.
- Hill, Sam, Richard Ettenson, and Dane Tyson (2005), "Achieving the Ideal Brand Portfolio," *MIT Sloan Management Review*, 46 (2), 85–90.
- Kolsarici, Ceren and Demetrios Vakratsas (2010), "Category- Versus Brand-Level Advertising Messages in a Highly Regulated Environment," *Journal of Marketing Research*, 47 (6), 1078–89.
- Kolsarici, Ceren and Demetrios Vakratsas (2015), "Correcting for Misspecification in Parameter Dynamics to Improve Forecast Accuracy with Adaptively Estimated Models," *Management Science*, 61 (10), 2495–513.
- Kruskal, William and Ruth Majors (1989), "Concepts of Relative Importance in Recent Scientific Literature," *The American Statistician*, 43 (1), 2–6.
- Lewis, Randall A. and Justin M. Rao (2015), "The Unfavorable Economics of Measuring the Returns to Advertising," *The Quarterly Journal of Economics*, 130 (4), 1941–73.
- Little, John D. C. (1966), "A Model of Adaptive Control of Promotional Spending," *Operations Research*, 14 (6), 1075–97.
- Little, John D. C. (1977), "Optimal Adaptive Control: A Multivariate Model for Marketing Applications," *IEEE Transactions on Automatic Control*, 22 (2), 187–95.
- Liu, Yan and Venkatesh Shankar (2015), "The Dynamic Impact of Product-Harm Crises on Brand Preference and Advertising Effectiveness: An Empirical Analysis of the Automobile Industry," *Management Science*, 61 (10), 2514–35.
- Low, George S. and Jakki J. Mohr (1999), "Setting Advertising and Promotion Budgets in Multi-Brand Companies," *Journal of Advertising Research*, 39 (1), 67–78.
- Low, George S. and Jakki J. Mohr (2000), "Advertising vs. Sales Promotion: A Brand Management Perspective," *Journal of Product & Brand Management*, 9 (6), 389–414.
- March, James G. and Herbert A. Simon (1958), *Organizations*. New York: Wiley.
- Naik, Prasad A., Murali K. Mantrala, and Alan G. Sawyer (1998), "Planning Media Schedules in the Presence of Dynamic Advertising Quality," *Marketing Science*, 17 (3), 214–35.
- Naik, Prasad A., Ashutosh Prasad, and Suresh Sethi (2008), "Building Brand Awareness in Dynamic Oligopoly Markets," *Management Science*, 54 (1), 129–38.
- Naik, Prasad and Kalyan Raman (2003), "Understanding the Impact of Synergy in Multimedia Communications," *Journal of Marketing Research*, 40 (4), 375–88.
- Naik, Prasad A., Kalyan Raman, and Russell S. Winer (2005), "Planning Marketing-Mix Strategies in the Presence of Interaction Effects," *Marketing Science*, 24 (1), 25–34.
- Nerlove, Marc and Kenneth J. Arrow (1962), "Optimal Advertising Policy Under Dynamic Conditions," *Economica*, 29 (114), 129–42.
- Osinga, Ernst C., Peter S. H. Leeflang, and Jaap E. Wieringa (2010), "Early Marketing Matters: A Time-Varying Parameter Approach to Persistence Modeling," *Journal of Marketing Research*, 47 (1), 173–85.
- Pekelman, Dov and Edison Tse (1980), "Experimentation and Budgeting in Advertising: An Adaptive Control Approach," *Operations Research*, 28 (2), 321–47.
- Piercy, Nigel F. (1987), "The Marketing Budgeting Process: Marketing Management Implications," *Journal of Marketing*, 51 (4), 45–59.
- Raman, Kalyan (1990), "Stochastically Optimal Advertising Policies Under Dynamic Conditions: The Ratio Rule," *Optimal Control Applications and Methods*, 11 (3), 283–88.
- Raman, Kalyan, Murali K. Mantrala, Shrihari Sridhar, and Yihui Elina Tang (2012), "Optimal Resource Allocation with Time-Varying Marketing Effectiveness, Margins, and Costs," *Journal of Interactive Marketing*, 26 (1), 43–52.
- Rosenzweig, Phil (2014), "The Benefits—and Limits—of Decision Models," *McKinsey Quarterly* (February), 1–10.
- Rubel, Olivier, Prasad A. Naik, and Shuba Srinivasan (2011), "Optimal Advertising When Envisioning a Product-Harm Crisis," *Marketing Science*, 30 (6), 1048–65.
- Rubel, Olivier and Ashutosh Prasad (2016), "Dynamic Incentives in Sales Force Compensation," *Marketing Science*, 35 (4), 676–89.
- Sahni, Navdeep S. (2015), "Effect of Temporal Spacing Between Advertising Exposures: Evidence from Online Field Experiments," *Quantitative Marketing and Economics*, 13 (3), 203–47.
- Sethi, Suresh P. and Gerald L. Thompson (2000), *Optimal Control Theory*, 2nd ed. New York: Springer.

- Sethuraman, Raj, Gerard J. Tellis, and Richard A. Briesch (2011), "How Well Does Advertising Work? Generalizations from Meta-Analysis of Brand Advertising Elasticities," *Journal of Marketing Research*, 48 (3), 457–71.
- Tull, Donald S., Van R. Wood, Dale Duhan, Tom Gillpatrick, Kim R. Robertson, and James G. Helgeson (1986), "'Leveraged' Decision Making in Advertising: The Flat Maximum Principle and Its Implications," *Journal of Marketing Research*, 23 (1), 25–32.
- Vakratsas, Demetrios and Prasad Naik (2007), "Essentials of Planning Media Schedules," in *Handbook of Advertising*, Tim Ambler and Gerard Tellis, eds. London: SAGE Publications.
- VanHeerde, Harald J., Kristiaan Helsen, and Marnik G. Dekimpe (2007), "The Impact of a Product-Harm Crisis on Marketing Effectiveness," *Marketing Science*, 26 (2), 230–45.
- Varian, Hal R. (2016), "Causal Inference in Economics and Marketing," *Proceedings of National Academy of Sciences*, 113 (27), 7310–15.
- Wang, Fang and Xiao-Ping Zhang (2008), "Reasons for Market Evolution and Budgeting Implications," *Journal of Marketing*, 72 (5), 15–30.
- West, Douglas, John B. Ford, and Paul W. Farris (2014), "How Corporate Cultures Drive Advertising and Promotion Budgets: Best Practices Combine Heuristics and Algorithmic Tools," *Journal of Advertising Research*, 54 (2), 149–62.
- White, Halbert (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48 (4), 817–38.
- Wittenmark, Björn (1975), "An Active Suboptimal Dual Controller for Systems with Stochastic Parameters," *Automatic Control Theory and Applications*, 3 (1), 13–19.