PRASAD A. NAIK and KALYAN RAMAN*

Many advertisers adopt the integrated marketing communications perspective that emphasizes the importance of synergy in planning multimedia activities. However, the role of synergy in multimedia communications is not well understood. Thus, the authors investigate the theoretical and empirical effects of synergy by extending a commonly used dynamic advertising model to multimedia environments. They illustrate how advertisers can estimate and infer the effectiveness of and synergy among multimedia communications by applying Kalman filtering methodology. Using market data on Dockers brand advertising, the authors first calibrate the extended model to establish the presence of synergy between television and print advertisements in consumer markets. Second, they derive theoretical propositions to understand the impact of synergy on media budget, media mix, and advertising carryover. One of the propositions reveals that as synergy increases, advertisers should not only increase the media budget but also allocate more funds to the less effective activity. The authors also discuss the implications for advertising overspending. Finally, the authors generalize the model to include multiple media, differential carryover, and asymmetrical synergy, and they identify topics for further research.

Understanding the Impact of Synergy in Multimedia Communications

Integrated marketing communications (IMC) emphasize the benefits of harnessing synergy across multiple media to build brand equity of products and services. Modern advertising textbooks adopt the IMC perspective (e.g., Belch and Belch 1998), major universities offer IMC courses (e.g., Petersen 1991), and many marketers and advertising agencies embrace the concept (e.g., *The New York Times* 1994). The American Association of Advertising Agencies (Schultz 1993, p. 17) defines IMC as follows: "A concept of marketing communications planning that recognizes the added value of [a] comprehensive plan that evaluates the strategic roles of a variety of communication disciplines—for example, general advertising, direct response, sales promotion, and public relations—and combines these disciplines to provide clarity, consistency, and maximum communications impact." This definition recognizes the added value aspect of IMC, which is created by the joint impact of multiple activities (e.g., television and print advertising). In other words, the combined effect of multiple activities exceeds the sum of their individual effects; this phenomenon is known as synergy (e.g., Belch and Belch 1998, p. 11).

Despite synergy's importance, its role in planning multimedia communications is not well understood (see Mantrala 2002). In a recent article, Schultz (2002, p. 6) observes that "consumers … live in a world of simultaneous media usage. They watch television while they surf the Net. They listen to radio while they read the newspaper. They page through a magazine while they download music from the Web.… What we really need today is a new approach to media planning, one that recognizes consumers' increasing ability to multitask and … [to] use a number of media simultaneously." Such an approach would elucidate the role of synergy in multimedia communications. For example, does synergy between television and print advertising exist in consumer markets? If synergy is present, how should brand managers measure it using readily available market data? Furthermore, how does its presence affect managers' decisions about size and allocation of the media budget? If synergy increases or decreases in a market, how should man-

*Prasad A. Naik is Associate Professor of Marketing, Graduate School of Management, University of California, Davis (e-mail: panaik@ucdavis.edu). Kalyan Raman is Professor of Marketing, School of Management, University of Michigan, Flint (e-mail: kalyanr@umflint.edu). The authors benefited from the suggestions of seminar participants at Columbia University; Erasmus University, Rotterdam; Harvard Business School; Katholieke Universiteit, Leuven; Massachusetts Institute of Technology; Penn State University; University of Chicago; University of Texas at Austin; and University of Texas at Dallas. For valuable comments and encouragement, the authors thank Eyal Biyalogorsky, Eitan Gerstner, Sunil Gupta, Gary Lilien, John D.C. Little, Vijay Mahajan, Kash Rangan, Arvind Rangaswamy, and Alvin Silk. Naik's research was supported in part by the University of California, Davis, Faculty Research Grant.

Journal of Marketing Research
Vol. XL (November 2003), 375–388

agers alter its media budget and the media mix? Finally, how does synergy moderate the effect of advertising carryover? The purpose of this article is to validate empirically and analyze theoretically the effects of synergy in multimedia communications.

To address these and related issues, we construct a model of multimedia communications. We extend a commonly used advertising model by incorporating the joint effects of multimedia communications. We calibrate the model by using proprietary advertising data from Levi Strauss that provide strong support for the presence of synergy between television and print advertising for the company's Dockers brand. We derive theoretical propositions to understand the impact of synergy on media budget, media mix, and advertising carryover. One of the propositions reveals that advertisers that experience synergy effects should increase the total budget and allocate more funds to the less effective activity. This counterintuitive result is a quintessential feature of the IMC framework.

We organize this article as follows: We review the extant literature to summarize previous work and differentiate our relative contributions. We then propose our IMC model, describe the data set, calibrate the model, perform model selection, check diagnostics, conduct cross-validation, and discuss empirical results. Next, we formulate the advertiser's budgeting and allocation problem and derive the optimal IMC strategy. Using the results, we investigate how advertisers should optimally respond to changes in market conditions given the presence of synergy. Subsequently, we generalize the model to the multiple media setting, allowing for differential carryover and asymmetrical synergy effects. Finally, we identify topics for further research and conclude by summarizing key IMC themes.

LITERATURE REVIEW

Interaction among marketing variables is a central theme in marketing. Indeed, it is interaction that provides a rigorous basis for the marketing-mix concept, "which emphasizes that marketing efforts create sales synergistically rather than independently" (Gatignon and Hanssens 1987, p. 247). Several studies document the joint effects of marketing variables on market outcomes. For example, advertising effectiveness increases with improved product quality (Kuehn 1962), greater retail availability (Parsons 1974), increased salesperson contact (Swinyard and Ray 1977), and a larger sales force (Gatignon and Hanssens 1987). Gatignon (1993) provides a comprehensive review of the literature on marketing interactions and describes the methods for calibrating models with interaction effects.

Notwithstanding this body of knowledge, and despite the fundamental significance of interactions in the marketing-mix concept, few studies systematically investigate the role of synergy in multimedia communications. For example, Sethi (1977) and Feichtinger, Hartl, and Sethi (1994, p. 219) comprehensively review the advertising control literature and conclude that "[w]ith a few exceptions, the [advertising] models assume … [a] single advertising medium. This was already noted by Sethi (1977), and this critical remark is still valid for the literature published subsequently." Among the few exceptions is Montgomery and Silk's (1972) study, which formalizes the concept of the communications mix; specifically, they define (p. B-485) communications mix as a "set of marketing activities by which a firm transmits product information and persuasive messages to a target market." Montgomery and Silk estimate the relative effectiveness of communications activities, such as product samples, direct mail, and television advertising for prescription drugs; however, they do not investigate the impact of synergy across these three media. Jagpal (1981) studied radio and print advertising for a commercial bank and was the first to present empirical evidence of synergy in multimedia advertising. However, his model ignores the carryover effect of advertising. Consequently, there is insufficient empirical evidence on the existence of cross-media synergy in dynamic markets. In addition, the literature contains no theoretical results on the effects of synergy on budgeting and allocation in dynamic markets.

Therefore, a consortium of radio network companies sponsored an industrywide field study, known as the "Image Transfer Study," to augment the sparse literature on cross-media synergy. Based on a sample of 500 adults, ages 20–44, from ten locations in Britain, the study indicated that 73% of participants remembered prime visual elements of television advertisements upon hearing radio commercials. In addition, 57% relived the television advertisements while listening to the radio advertisement. Thus, radio advertisements reinforced imagery created by television commercials, resulting in synergy between the two media. Information on randomized sampling and control procedures is not available because the industry research is proprietary.[1] More recently, Edell and Keller (1999) conducted controlled laboratory experiments and analyzed interactions between television and print advertisements to better understand the role of cross-media synergy.

A clear understanding of cross-media synergy is important, because it is likely to affect the allocation of marketing resources, as shown by Gatignon and Hanssens's (1987) and Gopalakrishna and Chatterjee's (1992) research in the area of personal selling. Specifically, these studies analyze the optimal resource allocation to advertising and sales force for a single-period (i.e., static) case; that is, their normative analysis ignores the advertising carryover effect. In addition, they do not investigate how optimal allocation varies with the magnitude of synergy. Table 1 further identifies the main differences and relative contributions of different studies. Thus, the problem of optimal budgeting and allocation in dynamic markets in the presence of synergy among communications activities remains unsolved.

In summary, the literature provides limited empirical and theoretical knowledge on cross-media synergy. Thus, it offers little guidance to managers about optimal allocation of resources across multiple media in dynamic markets. To this end, we formulate an IMC model.

MODEL DEVELOPMENT

We begin with a simple first-order autoregressive advertising model (e.g., Palda 1964), which is commonly used in practice (Bucklin and Gupta 1999, p. 262):

$$S_t = \alpha + \beta u_t + \lambda S_{t-1} + \nu_t, \tag{1}$$

where S_t is sales at time t, u_t is advertising effort at time t, α represents the mean level of initial sales in the absence of

[1]For additional information, contact the Radio Advertising Bureau or visit http://www.rab.co.uk.

Table 1
MAIN DIFFERENCES AMONG RELATED STUDIES

Features	*Gatignon and Hanssens (1987)*	*Gopalakrishna and Chatterjee (1992)*	*Present Study*
Sales response model	Static	Dynamic	Dynamic
Interactions	Sales force by advertising	Sales force by advertising	Television by print advertising
Estimation method	Generalized least squares regression	Nonlinear least squares regression	Kalman filter estimation
Model selection	No	No	Yes
Cross-validation	No	No	Yes
Tests exogeneity	No	No	Yes
Characterizes optimal decisions	Numerically, for the static case	Analytically, for the static case	Analytically, for the dynamic case
Derives comparative statics	No	Yes, static case	Yes, dynamic case
Investigates effects of interaction on optimal decisions	No	No	Yes
N-media generalization	No	No	Yes
Remarks	Proposes process functions for static parameters and develops an estimation approach to calibrate them.	—	Extends process functions to dynamic, nonstationary, random parameters and develops an estimation approach to calibrate them.

advertising (i.e., $\alpha = E[S_0]|u_{t<0} = 0$), β is the short-term effect of advertising, λ is the carryover effect of advertising, and ν_t is a normally distributed error term that represents the impact of other factors that are not explicitly included in the model for the sake of parsimony.

Montgomery and Silk (1972) extend Equation 1 to incorporate the effects of multimedia advertising by conceptualizing advertising as a mix of multimedia activities, each with different effectiveness. Consequently, the impact of two activities $(u, v)'$ with unequal effectiveness parameters $(\beta_1, \beta_2)'$ is given by

$$S_t = \alpha + \beta_1 u_t + \beta_2 v_t + \lambda S_{t-1} + \nu_t. \tag{2}$$

Finally, following Jagpal (1981) and Gopalakrishna and Chatterjee (1992), we introduce an interaction term to capture the joint effects of multimedia activities:

$$S_t = \alpha + \beta_1 u_t + \beta_2 v_t + \kappa u_t v_t + \lambda S_{t-1} + \nu_t. \tag{3}$$

The conceptual distinction between Equations 2 and 3 is the following: In Equation 2, both advertising media (u, v) increase brand sales S, because β_1 and β_2 are expected to be positive. In Equation 3, advertising serves a dual purpose: It increases sales and enhances media effectiveness. If $\kappa > 0$, advertising increases the effectiveness of the other medium. Thus, Equation 3 introduces the role of synergy because the combined sales impact of (u, v) exceeds the sum of the independent effects $(\beta_1 u + \beta_2 v)$ when $\kappa > 0$. The subsequent remarks elaborate on the model formulation.

Remark 1

Two different mechanisms can give rise to Equation 1. According to the Koyck model (see Hanssens, Parsons, and Schultz 1998, p. 215), the short-term effect of advertising on sales is β, so that the subsequent period effects are $\lambda\beta$, $\lambda^2\beta$, $\lambda^3\beta$, and so on. The long-term impact of advertising is obtained from the sum $\sum_{i=0}^{\infty}\lambda^i\beta = \beta/(1 - \lambda)$. According to the partial adjustment model (see Leeflang et al. 2000, p. 95), managers attain the target level of sales, $S_t^{\#}$, by advertising so that $S_t^{\#} = \alpha^{\#} + \beta^{\#}u_t + \varepsilon_t^{\#}$. After consumers adjust to such managerial actions, the market grows by $S_t - S_{t-1} = (1 - \lambda)(S_t^{\#} - S_{t-1})$. Thus, managerial actions drive market growth over time, leading to Equation 1, which can be obtained by eliminating $S_t^{\#}$. These two mechanisms imply different error structures, which can be distinguished by examining the properties of residuals. Rejection of serial correlation in the estimated residuals lends support for the partial adjustment mechanism.

Remark 2

Gatignon and Hanssens (1987) propose the process function view, which provides a rationale for Equation 3. The main idea is that managerial actions affect not only market outcomes but also the effectiveness of marketing activities. For example, suppose that radio advertising increases sales and enhances television advertising effectiveness. These effects are captured in the process function $\beta_1 = \beta_1' + \kappa v_t$, which on substitution in Equation 2 leads to Equation 3. We note that this process function has no error term. In the "Discussion" section, we specify a broader class of response and process functions that involve both stochastic and dynamic components, and we outline an approach for their estimation and inference.

Remark 3

On the basis of the Image Transfer Study and Edell and Keller's (1999) laboratory experiments, we note that the mutually reinforcing effects of various media create synergy. For example, when consumers attend to radio or print advertisements, they remember the images they have viewed in previous television advertisements. Similarly, billboards and in-transit advertisements remind consumers of messages or images from radio and television. This process strengthens brand knowledge in their memory, thus enhancing the combined impact of multiple media (Keller 1998, p. 257).

Remark 4

We investigate the possibility of relaxing the constant carryover assumption in the simplest multimedia model.[2] Consider the extension of Equation 2 in which carryover effects are unequal:

$$(4) \qquad S_t = \alpha + \beta_1 u_t + \beta_2 v_t + \lambda_1 S^1_{t-1} + \lambda_2 S^2_{t-1} + \nu_t,$$

where S^1 and S^2 denote sales due to television and print, respectively, so that (λ_1, λ_2) are carryover effects of television and print media. Note that because brand sales are not available by each medium, the coefficients (λ_1, λ_2) cannot be estimated separately. This lack of information creates an "identification problem" (see Rothenberg 1971). One way to overcome this problem is to identify conditions that supplement the unavailable information set. To achieve this, we let $S^i_t = w_i S_t$ and i = 1, 2, where w_i is the proportion of the total sales due to medium i, and so $\Sigma_i w_i = 1$. Then, substituting S^i in Equation 4, we obtain

$$(5) \qquad S_t = \alpha + \beta_1 u_t + \beta_2 v_t + [\lambda_1 w_1 + \lambda_2(1 - w_1)]S_{t-1} + \nu_t.$$

Equation 5 eliminates the dependence on the unknown sales S^1 and S^2 and uses known information on the lagged sales S_{t-1}. We need to assume that $\lambda_2 = k\lambda_1$ to enable the estimation of the coefficient of lagged sales. Then, the carryover effect is due to the television medium when k = 0, the carryover effect is attributable to both media when k = 1, and we obtain unequal media-specific carryover effects for intermediate values of k. However, we must know w_1 to estimate the lagged-sales coefficient, $\lambda_1[w_1 + k(1 - w_1)]$, by prespecifying the values of k. In other words, we must either know w_i (which is the same as knowing S^i) or supplement an identification condition $w_i = w$. For Equation 5, under this condition, the lagged-sales coefficient becomes $\lambda_1(1 + k)/2$, which can be estimated as $\hat{\lambda}_1(k)$ for a given value of k. The best k^* is the one that maximizes the log-likelihood function, which is found through grid search, thus yielding the media-specific carryover effects $\hat{\lambda}^*_1 = \hat{\lambda}_1(k)$ and $\hat{\lambda}_2 = k^*\hat{\lambda}^*_1$. Thus, without these assumptions, media-specific carryover effects are inestimable even for the simplest multimedia model (e.g., Zellner and Geisel 1970).

In summary, we must assume that either the carryover effect is the same across media or the proportion of sales attributable to either medium is the same, though neither assumption is entirely realistic. We illustrate this approach subsequently (see Equation 9) and note that the substantive conclusions in terms of synergy effects hold irrespective of the carryover effect being constant or different across media.

Remark 5

A purpose of advertising is to increase the retention rate of customers. To this end, we consider the following extension of Equation 3:

$$(6) \qquad S_t = \alpha + \beta_1 u_t + \beta_2 v_t + \lambda S_{t-1} + \kappa_1 u_t v_t + \kappa_2 u_t S_{t-1} + \kappa_3 v_t S_{t-1} + \nu_t,$$

which includes the interactions between advertising media and lagged sales.[3] In Equation 6, the parameters (κ_2, κ_3) represent the moderating effects of current advertising on carryover effect; that is, managers can determine whether current advertising amplifies or attenuates the carryover effect. We subsequently estimate and compare this model.

Remark 6

Despite the deceptively simple appearance of Equations 3 and 6, they are stochastic difference equations with nonlinearity in the decision variables. Consequently, they induce intertemporal dependence between sales observations, which must be incorporated in parameter estimation. Because the ordinary least squares (OLS) approach uses the marginal density of sales to construct the likelihood function, it ignores this intertemporal dependence. Therefore, OLS estimates are biased (for Monte Carlo evidence, see Naik and Tsai 2000a). In addition, OLS estimation yields inconsistent estimates unless the error term is serially uncorrelated, an assumption that may not be satisfied in practical applications. Thus, we apply Kalman filter estimation to calibrate dynamic models with real data, which we describe next for Dockers brand advertising.[4]

EMPIRICAL ANALYSES

In this section, we describe the data set and the estimation method. We then present the empirical results, test alternative specifications, check diagnostics, and cross-validate the model.

Data Description

We study Levi Strauss's advertising decisions for Dockers khaki pants fashion apparel. The company's advertising agency, Foote, Cone & Belding, developed the "Nice Pants" advertising campaign. An execution of the campaign shows a young man noticing a beautiful woman on a subway train. He tries to reach out to her, but the train doors close. As the train departs, he sees her mime the compliment "Nice pants." The advertisements do not show the advertised product, and thus the creative execution is considered offbeat (Enrico 1996). However, the brand managers at Levi Strauss deemed the campaign a success because it increased sales.

The market data consist of retail sales in thousands of units sold and expenditures on network television and print advertising from 1994 to 1997. To maintain confidentiality, we rescale this proprietary data.[5] We denote retail sales with S_t, network advertising with u_t, and print advertising with v_t, t = 1, …, 47 months. Figure 1 displays the actual sales in

[2]We thank a reviewer for suggesting this analysis.

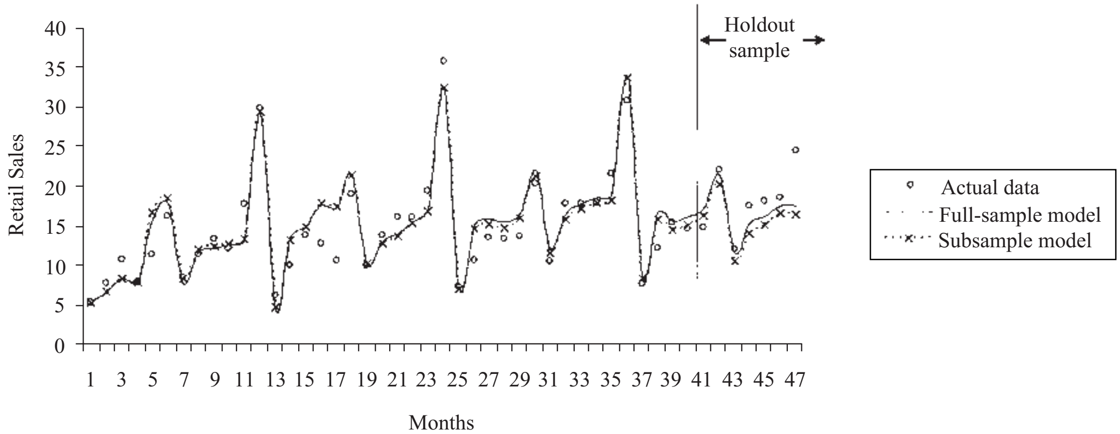
[3]We thank a reviewer for suggesting this model.

[4]The classic references on OLS and Kalman filtering are Rao (1973) and Jazwinski (1970), respectively. For small sample properties of Kalman filtering, see Oud, Jansen, and Haughton (1999).

[5]Despite that Levi Strauss is a private enterprise, we acknowledge that the company's executives were not only cooperative but also progressive in their attitude toward adopting quantitative methods of marketing science.

Product and media prices were fairly constant during the period of analysis. Specifically, consumer and producer prices for men's apparel exhibit an annual increase of .98% and .54%, respectively. Similarly, the producer price index for national advertising shows a gradual increase of 5% per annum. We note that the reestimated model using deflated dollar advertising expenditures leads to similar substantive conclusions.

Figure 1
ACTUAL RETAIL SALES VERSUS MODEL FORECASTS



which we observe seasonal effects at the end of the calendar year (during the Christmas selling season) and a small rise and fall in June and July, respectively. We construct dummy variables to capture the seasonal and midyear effects as follows:[6]

$$D_{1t} = \begin{cases} 1 & \text{if } t = 12, 24, ..., N \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

$$D_{2t} = \begin{cases} 1 & \text{if } t = 13, 25, ..., N \\ 0 & \text{otherwise.} \end{cases}$$

$$D_{3t} = \begin{cases} 1 & \text{if } t = 6, 18, ..., N \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

$$D_{4t} = \begin{cases} 1 & \text{if } t = 7, 19, ..., N \\ 0 & \text{otherwise.} \end{cases}$$

Next, we apply Kalman filter estimation to these data to calibrate the model (Equation 3).

Kalman Filter Estimation

We design a Kalman filter by obtaining a transition equation based on the model dynamics and linking it to observed sales by means of an observation equation. Typically, the transition equation includes factors that influence the dynamics (i.e., change in sales), whereas the observation equation incorporates factors such as seasonality that affect the level of observed sales. Using transition and observation equations, we compute the likelihood of observing a sales trajectory as the product of conditional densities, given the history of sales and advertising. Finally, we estimate the parameters and their standard errors by applying the standard principles of maximum-likelihood estimation (for further details, see Naik, Mantrala, and Sawyer 1998).

We obtain the transition equation by decomposing the total sales into three components: (1) sales due to television advertising, S_{1t}; (2) sales due to print advertising, S_{2t}; and (3) sales due to synergy between television and print advertising, S_{3t}. Thus, we reexpress Equation 3 in the following vector form:

$$\begin{bmatrix} S_{1,t} \\ S_{2,t} \\ S_{3,t} \end{bmatrix} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \begin{bmatrix} S_{1,t-1} \\ S_{2,t-1} \\ S_{3,t-1} \end{bmatrix} + \begin{bmatrix} \beta_1 u_t \\ \beta_2 v_t \\ \kappa u_t v_t \end{bmatrix} + \begin{bmatrix} \nu_{1t} \\ \nu_{2t} \\ \nu_{3t} \end{bmatrix}, \quad (9)$$

where $\nu_{it} \sim N(0, \sigma^2/3)$ for i = 1, 2, and 3.[7] By summing both sides of Equation 9 and noting that the total sales $S_{t-\tau} = S_{1,t-\tau} + S_{2,t-\tau} + S_{3,t-\tau}$, where $\tau = 0, 1$, we obtain Equation 3. We introduce differential carryover in Equation 9 and then diminishing returns and seasonal effects.

The carryover effect is the constant λ in the principal diagonal of the transition matrix (denoted by T) on the right-hand side of Equation 9. Suppose that print carryover effect is not equal to overall carryover effect. To explore this, we specify the second diagonal element in the transition matrix as $k\lambda$; that is, the transition matrix $T = \text{diag}(\lambda, k\lambda, \lambda)$ in Equation 9. As we explained in Remark 4, to estimate differential carryover effects, we apply the identification condition $w_i = w$, where $w_i = S_{i,t-\tau}/S_t$, for determining the best value of k through grid search. Thus, managers can assess the possibility of unequal carryover effect for print advertising.[8]

Next, to incorporate diminishing returns, we operationalize $u_t = \sqrt{x_{1t}}$ and $v_t = \sqrt{x_{2t}}$, where x_{1t} and x_{2t} denote television and print advertising expenditures, respectively. We then link Equation 9 to observed sales Y_t, which includes the seasonality and midyear effects $\gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)'$:

$$Y_t = [1 \quad 1 \quad 1] \begin{bmatrix} S_{1,t} \\ S_{2,t} \\ S_{3,t} \end{bmatrix} + \sum_{i=1}^{4} \gamma_i D_{it} + \varepsilon_t, \quad (10)$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$.

In Equations 9 and 10, the two errors $(\nu_t', \varepsilon_t)'$ conceptually represent different kinds of uncertainty: One is inherent in modeling a dynamic system, and the other arises in measuring an observed system. Substantively, transition uncertainty

[6]We acknowledge that this specification implies fixed seasonal effects.

[7]We test that error terms satisfy the set of assumptions (see "Model Diagnostics"). Their violations indicate that some signal in the data can be extracted by the filter.

[8]We thank a reviewer for suggesting this grid search approach.

relates to the understanding (or the lack thereof) of how advertising dynamics work for a given product market; observation uncertainty is induced by the process of measurement, such as the use of accounting measures or tracking surveys. Mathematically, the error terms $(\nu_t', \varepsilon_t)'$ follow the multivariate normal distribution:

$$(11) \quad N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & P' \\ P & H \end{bmatrix}\right).$$

The two kinds of errors are orthogonal to each other (i.e., independent) because they are qualitatively different and possess different meanings, and thus the default value of P = 0 (e.g., see Hamilton 1994, Ch. 13). More specifically, we have $Q_{3 \times 3} = \text{diag}(\sigma^2/3, \sigma^2/3, \sigma^2/3)$, $H_{1 \times 1} = \sigma_\varepsilon^2$, and $P_{1 \times 3} = (0, 0, 0)$ in this filter design.

Using the transition and observation Equations 9 and 10, we compute the log-likelihood of observing the sales trajectory $Y = (Y_1, Y_2, \ldots, Y_T)'$, which is given by

$$(12) \quad LL(\Theta; Y) = \sum_{t=1}^{T} g\left(Y_t \middle| \Im_{t-1}\right),$$

where $g(\cdot \mid \cdot)$ is the conditional density of sales Y_t, given the history to the last period, $\Im_{t-1}$. The random variable $Y_t|\Im_{t-1}$ is normally distributed for all t, and its mean and variance are given by a set of difference equations known as the Kalman filter (see Naik, Mantrala, and Sawyer 1998, p. 234). The vector Θ contains model parameters $(\lambda, \beta_1, \beta_2, \kappa, \gamma_1, \gamma_2, \gamma_3, \gamma_4)'$ and other parameters such as variances of error terms in transition and observation equations and initial means of the state vector. By maximizing Equation 12 with respect to Θ, we obtain the maximum-likelihood Kalman filter estimates, $\hat{\Theta}$. The standard errors of parameter estimates are obtained from the information matrix evaluated at estimated values. These estimates are asymptotically unbiased and possess minimum variance among all estimators because the transition and observation equations are linear in the state variables S_{it} and the error terms $(\nu_{it}, \varepsilon_t)'$ are normally distributed (Harvey 1994, p. 110).

Estimation Results

We estimate the empirical analog of Equation 3 given by the extended Equations 9 and 10, which include differential carryover, diminishing returns, and seasonal effects. Table 2 presents the parameter estimates, standard errors, and t-values. The t-values indicate that all estimates are statistically significant at the 95% confidence level (except the print advertising effectiveness). The carryover effect of television advertising is $\hat{\lambda} = .9269$. When we performed the grid search over different values of k = 0 (.1)1, we found that k = .4 results in the largest value of LL*(k), which suggests that print carryover is approximately 40% of the overall carryover effect.[9] The effectiveness of television advertising is large and significant, whereas that of print advertising is not significant. More important, the synergy between television and print advertising is large ($\hat{\kappa} = 1.5766$) and significant (t-value = 2.4). In addition, we find that retail sales increase by $\hat{\gamma}_1 = 15.3877 \times 100{,}000$ units during Christmas and drop by $\hat{\gamma}_2 = 9.3741 \times 100{,}000$ units in the following month, presumably as a result of stockpiling and product returns by consumers. Similarly, the company should expect sales to increase by $\hat{\gamma}_3 = 5.4238 \times 100{,}000$ units in June and to decrease by $\hat{\gamma}_4 = 3.1974 \times 100{,}000$ units in July. Figure 1 presents the model forecasts and actual sales, indicating a good fit.

Model Selection, Diagnostics, and Cross-Validation

Model selection. The goal of model selection is to select the most parsimonious model supported by the observed data. To balance parsimony (retain few parameters) and fidelity (enhance goodness-of-fit), we compute three information criteria: Akaike information criterion (AIC), its bias-corrected version (AIC_C; Hurvich and Tsai 1989), and Schwarz's information criterion (BIC).[10] We select a model associated with the smallest values of the information criteria. Table 3 presents the values of the information criteria for all models, with and without synergy, and models with other interaction terms. Specifically, the proposed model attains the AIC value of 149.29, which is the smallest compared with other specifications. Similarly, its AIC_C value is 207.47, which is less than AIC_C values for other models. Furthermore, the minimum BIC value of 169.65 is attained by the proposed model. Thus, all three information criteria

[9] If an interior solution is not found, grid search must be extended beyond the unit interval.

[10] For computation, $AIC = -2LL^* + 2K$, $AIC_C = -2LL^* + T(T + K)/(T - K - 2)$, and $BIC = -2LL + KLn(T)$, where LL^* is the maximized log-likelihood value, T is the sample size, and K is the number of parameters.

Table 2
KALMAN FILTER ESTIMATES

Model Parameters	*Estimates*	*Standard Errors*	*t-Values*
Carryover effect, λ	.9269	.0313	29.62
Print carryover effect = $k\lambda$, where k = .4[a]	.3708	—	—
Effectiveness of television advertisements, β_1	1.3530	.5536	2.44
Effectiveness of print advertisements, β_2	2.1724	1.1824	1.84
Synergy between television and print, κ	1.5766	.6570	2.40
Christmas buying spree, γ_1	15.3877	1.4528	10.59
Post-Christmas drop, γ_2	–9.3741	1.3914	–6.74
Midyear effect, γ_3	5.4238	1.2066	4.49
Post-midyear effect, γ_4	–3.1974	1.3614	–2.35
Model fit, R^2		82.51%	
Maximized log-likelihood value, LL^*		–63.65	
AIC		149.29	
AIC_C		207.47	

[a] Note that k = .4 results in the largest value of LL*(k).

Table 3
VALUES FOR INFORMATION CRITERIA

Models[a]	*AIC*	*AIC$_C$*	*BIC*
Equation 2, no synergy	153.12	209.66	171.62
Equation 3, synergy	149.29	207.47	169.65
Equation 6, only television by lagged sales interaction	149.93	209.96	172.12
Equation 6, only print by lagged sales interaction	151.22	211.25	173.42
Equation 6, both television and print interactions with lagged sales	151.93	214.05	175.97

[a]All estimated empirical analogs of these models include differential carryover and diminishing returns.

support the retention of the proposed model, providing strong convergent validity and enhancing our confidence in the retained model.

Model diagnostics. We test the model residuals for skewness, kurtosis, heteroskedasticity, serial correlation, model stability, and multicollinearity. Because these tests are standard (e.g., Harvey 1994, p. 256–60), we simply report the findings: Residuals are normally distributed, they do not exhibit either heteroskedasticity or serial correlation, the estimated parameters are stable over time, and multicollinearity is not present. The lack of serial correlation suggests that the model dynamics are likely to be driven by the partial adjustment mechanism (see "Remark 1"). Furthermore, on the basis of the likelihood ratio test, we retain a common variance for the three error terms in Equation 9.

Using Engle, Hendry, and Richard's (1983) approach, we also test for the exogeneity of advertising. We find that both television and print advertising not only are weakly exogenous but also satisfy strong and superexogeneity requirements. Thus, the estimated Equation 9 is adequate for efficient estimation (because of weak exogeneity), forecasting (because of strong exogeneity), and policy simulation (because of superexogeneity).

Cross-validation. To test predictive accuracy on out-of-sample data, we conduct a cross-validation study. Specifically, using 40 observations, we calibrate Equation 9 with differential carryover and diminishing returns, and we forecast the remaining 7 observations in the holdout sample. The correlation between the holdout observations and their forecast from the full sample calibrated model is .738, whereas that from the subsample calibrated model is .736. That is, compared with the correlation for the full-sample model, which sets an upper bound because it uses all information from the sample, the correlation for the subsample model drops only marginally. Thus, relative to the full-sample model, the subsample model performs satisfactorily when used to predict out-of-sample observations. Figure 1 displays the cross-validation results and shows how the subsample model conservatively predicts holdout observations and correctly anticipates turning points.

In summary, the statistical analyses furnish strong evidence for the presence of synergy between television and print advertising for Dockers brand. Note that brand managers can implement the Kalman filter approach to estimate and infer the existence of synergy using readily available market data on sales and advertising decisions.

NORMATIVE ANALYSES

Given the presence of synergy, how should brand managers determine the media budget? How should they alter the media mix if synergy increases in some markets? We present normative analyses to address such substantive issues. We formulate an advertiser's decision-making problem and then derive the optimal IMC strategy.

Advertiser's Decision-Making Problem

An advertiser's decision-making problem is to determine the total budget and its allocation to various communication activities. Suppose that the advertiser decides to expend effort on two activities over time as follows: $\{u_1, u_2, \ldots, u_t, \ldots\}$ and $\{v_1, v_2, \ldots, v_t, \ldots\}$. Given this specific media plan, $\{(u_t, v_t) : t \in (1, 2, \ldots)\}$, the advertiser generates the sales sequence $\{S_1, S_2, \ldots, S_t, \ldots\}$ and earns an associated stream of profits $\{\pi_1, \pi_2, \ldots, \pi_t, \ldots\}$. Discounting future profits at the rate ρ, the advertiser computes the net present value $J = \Sigma_{t=1}^{\infty} e^{-\rho t}\pi_t(S_t,u_t,v_t)$. Thus, a media plan $(u, v) = \{(u_t, v_t) : t = 1, 2, \ldots\}$ induces a sequence of sales that yields a stream of profits with a net present value of $J(u, v)$. Formally, the budgeting problem (see Little 1979) is to find the optimal plan (u^*, v^*) that attains the maximum value J^*.

An advertiser could solve this budgeting problem by using the following algorithm: (1) Specify a trial plan (u, v); (2) use the parameter estimates in Table 2 and develop the sales forecast under the trial plan; (3) compute $\pi(S_t, u_t, v_t)$ and J(u, v), making reasonable assumptions; (4) select a new plan (u, v) to increase J; and (5) continue Steps 2–4 until J^* is attained. This numerical algorithm can determine a nearly optimal sequence of advertising effort levels for specific brands, such as Dockers; however, it will not yield generalizable insights into how advertisers, facing any set of feasible parameter values, should behave optimally. In other words, how should advertisers alter the media budget ($B_t = u_t^* + v_t^*$) and media mix ($\Lambda_t = u_t^*/v_t^*$) as market conditions change? Generalizable insights may be obtained by applying deterministic optimal control theory (e.g., Kamien and Schwartz 1991) to the continuous-time version of the preceding problem.

Denote the optimal plan for the two activities by $u^*(t)$ and $v^*(t)$. Formally, the advertiser seeks to determine $u^*(t)$ and $v^*(t)$ by maximizing

$$J(u, v) = \int_0^{\infty} e^{-\rho t}[S(t), u(t), v(t)]dt, \tag{13}$$

where ρ denotes the discount rate, $\Pi(S, u, v) = mS - u^2 - v^2$ is the profit function (because $u = \sqrt{x_1}$ and $v = \sqrt{x_2}$, where $[x_1, x_2]$ are media expenditures), m is unit profit margin, and J(u, v) is the net present value of any multimedia policies (u[t], v[t]).

In finding the optimal IMC strategy $\{[u^*(t), v^*(t)] : t \in [0,\infty)\}$, the advertiser needs to account for the impact of synergy and the dynamics of advertising response. The roles of synergy and dynamics are embodied in the following state equation:

$$\frac{dS}{dt} = \lim_{\Delta t \to 0} \frac{\Delta S_t}{\Delta t} = \beta_1 u(t) + \beta_2 v(t) + \kappa u(t)v(t) - (1 - \lambda)S(t), \tag{14}$$

which is the deterministic continuous-time version of Equation 3. When Equation 14 explicitly includes parametric or environmental uncertainty, more advanced analysis using stochastic control theory is required (see the "Discussion" section).

Optimal IMC Strategy

We solve the maximization problem induced by Equations 13 and 14 by applying optimal control theory. We present the derivation of the optimal strategies in Appendix A to maintain continuity of exposition, and we state the results of our optimality analysis here. We find that the optimal level of the first communications activity is given by

$$u^* = \frac{m[\beta_2 \kappa m + 2\beta_1(1 + \rho - \lambda)]}{4(1 + \rho - \lambda)^2 - \kappa^2 m^2}, \tag{15}$$

and the optimal level for the second activity is

$$v^* = \frac{m[\beta_1 \kappa m + 2\beta_2(1 + \rho - \lambda)]}{4(1 + \rho - \lambda)^2 - \kappa^2 m^2}. \tag{16}$$

Using Equations 15 and 16, we obtain the total budget $B = u^* + v^*$ as

$$B = \frac{(\beta_1 + \beta_2)m}{2(1 + \rho - \lambda) - \kappa m}. \tag{17}$$

Similarly, we get the optimal media mix $\Lambda = u^*/v^*$ as

$$\Lambda = \frac{2\beta_1(1 + \rho - \lambda) + m\beta_2\kappa}{2\beta_2(1 + \rho - \lambda) + m\beta_1\kappa}. \tag{18}$$

In practice, advertisers should estimate the model parameters for their specific brand in a given market by applying the Kalman filter estimation approach we explained previously and then determine the optimal budget and media mix from Equations 17 and 18, respectively. We use these equations to generate theoretical insights into optimal budgeting and allocation in the presence of synergy and discuss these findings in the next section.

MARKETING IMPLICATIONS

In this section, we augment the extant advertising literature with new propositions that elucidate the role of synergy in multimedia communications. We derive the propositions by means of comparative dynamics analysis of the equilibrium strategies (see Kamien and Schwartz 1991).[11] For expositional clarity, we relegate all the proofs to Appendix B and focus on the results, intuition, and implications.

Multimedia Budgeting and Allocation in the Absence of Synergy

Dorfman and Steiner's (1954) results provide a meaningful background for the new propositions. Using a static advertising model, they show that the optimal advertising-to-sales ratio is proportional to the ratio of the advertising and price elasticities. Thus, if advertising effectiveness improves, advertisers should increase the advertising budget, ceteris paribus. Nerlove and Arrow (1962) demonstrate that this result holds even in the presence of dynamics arising from the carryover effect. We note that both studies investigate optimal spending on a single advertising medium.[12] In the multimedia context, two conceptual issues arise: First, why should advertisers spend anything at all on the second, less effective medium? Second, if the effectiveness of one medium were to increase, should they reallocate the same total budget in proportion to the new effectiveness or increase the total budget as well? The following proposition addresses these issues.

P_1: In multimedia advertising, as the effectiveness of an activity increases, the advertiser should increase spending on that activity, thus increasing the media budget. Furthermore, the media budget should be allocated to various activities in proportion to their relative effectiveness.

Multimedia Budgeting in the Presence of Synergy

P_2: As synergy increases, the advertiser should increase the media budget.

This result illuminates the issue of overspending in advertising. The marketing literature (see Hanssens, Parsons, and Schultz 1998, p. 260) suggests that advertisers overspend; that is, the actual expenditure exceeds the optimal budget implied by normative models. However, the claim that advertisers overspend is likely to be overstated in the IMC context, because the optimal budget itself is understated when response models ignore the impact of synergy. To verify, we compute the optimal budget from Equation 17 with synergy ($\kappa \neq 0$) and without it ($\kappa = 0$). We then find that the optimal budget required for managing multimedia activities in the presence of synergy is greater than that required in its absence. Thus, in practice, if the advertiser's budget reflects the objective of integrating multimedia communications, overspending is likely to be smaller. Further research may extend these findings when parameters are uncertain and firms are risk averse. Next, we show that budget allocation between activities is qualitatively different in the presence of synergy.

Multimedia Allocation in the Presence of Synergy

P_3: As synergy increases, the advertiser should decrease (increase) the proportion of media budget allocated to the more (less) effective communications activity. If the various activities are equally effective, the advertiser should allocate the media budget equally among them, regardless of the magnitude of synergy.

Because this result is counterintuitive, we provide a detailed explanation. First, observe that the first-order conditions establish the following relationship between the optimal spending levels: $u^* \propto \beta_1 + \kappa v^*$ and $v^* \propto \beta_2 + \kappa u^*$. Second, suppose that two activities have unequal effectiveness (e.g., $\beta_1 > \beta_2$). In the absence of synergy ($\kappa = 0$), the

[11]Comparative dynamics involve analysis of a change in the dynamic equilibrium strategy with respect to a change in the model parameter. Kamien and Schwartz (1991, p. 168) note that in contrast to comparative statics, this analysis is "more difficult, but sometimes possible."

[12]Although we use monetary terms such as "spending," the models provide general insights for expending effort to allocate resources other than just dollars (e.g., time, attention).

optimal spending on an activity depends only on its own effectiveness; thus, a larger amount is allocated to the more effective activity (see P_1). However, in the presence of synergy ($\kappa \neq 0$), optimal spending depends not only on its own effectiveness but also on the spending level for the other activity. Consequently, as synergy increases, marginal spending on an activity increases at a rate proportional to the spending level for the other activity (i.e., $\partial u^*/\partial\kappa \propto v^*$ and $\partial v^*/\partial\kappa \propto u^*$). Thus, optimal spending on the more effective activity increases slowly, relative to the increase in the optimal spending on the less effective activity (i.e., $\partial u^*/\partial\kappa < \partial v^*/\partial\kappa$ because $v^* < u^*$). Thus, the proportion of budget allocated to the more effective activity decreases as synergy increases.

If the two activities are equally effective, the optimal spending levels for both are equal. Furthermore, as synergy increases, marginal spending on each activity increases at the same rate. Thus, the optimal allocation ratio remains constant at 50%, regardless of the increase or decrease in synergy.

Another notable perspective is as follows: The optimal budget is always more heavily allocated toward the more effective activity.[13] For a fixed budget ($u_t + v_t$), the impact of the interaction term is greatest when $u_t = v_t$. Thus, as κ increases (holding β_1 and β_2 constant), the ratio of the two expenditures should become less lopsided in favor of the more effective activity (i.e., Λ moves toward 1 as κ increases). We next study how synergy moderates the carryover effect.

Advertising Carryover Effect in the Presence of Synergy

P_4 (budget): As the carryover effect increases, the advertiser should increase the media budget. This rate of increase in the media budget increases as synergy increases.

P_5 (allocation): In the absence of synergy, budget allocation does not depend on the carryover effect; in contrast, it depends on the carryover effect in the presence of synergy. In the latter case, as carryover increases (decreases), the advertiser should decrease (increase) the proportion of budget allocated to the more (less) effective activity.

P_4 and P_5 show that advertisers should allocate their budgets differently in markets with and without synergy. In the absence of synergy, advertisers should allocate the budget to various activities in simple proportion to their relative effectiveness; in markets with synergy, the allocation should take into account the magnitude of the carryover effect. In the next section, we extend the model to the N-media setting.

N-MEDIA GENERALIZATION WITH DIFFERENTIAL CARRYOVER AND ASYMMETRIC SYNERGY

We further extend the IMC model to the N-media setting and incorporate the effects of unequal carryover and asymmetrical synergy. Using two illustrative scenarios, we show the application of the extended IMC theory to generate new insights systematically.

Let β_{12} denote the synergy between Medium 1 and Medium 2 ($u_1 \to u_2$), and let β_{21} be the synergy between Medium 2 and Medium 1 ($u_2 \to u_1$). For example, broadcast advertising enhances sales force effectiveness, but sales force communications may not increase advertising effectiveness to the same extent.

An advertiser may use N different media, which we denote by the vector $\vec{u} = (u_1, ..., u_N)'$. Then, the objective function corresponding to Equation 13 is

$$J(u_1, ..., u_N) = \int_0^\infty e^{-\rho t}\Pi[S(t), \vec{u}(t)]dt, \tag{19}$$

where $\Pi(S, \vec{u}) = mS - \Sigma_{i=1}^N u_i^2$, $S = \Sigma_{i=1}^N S_i$, and S_i denotes sales generated by medium i, i = 1, …, N. Equation 14 generalizes to

$$\frac{dS}{dt} = \sum_{i=1}^{N}\frac{dS_i}{dt}, \tag{20}$$

where each S_i evolves according to

$$\frac{dS_i}{dt} = \beta_i u_i + \sum_{\substack{j=1\\ j\neq i}}^{N}\beta_{ij}u_i u_j - (1-\lambda_i)S_i. \tag{21}$$

In Equations 20 and 21, we introduce asymmetrical synergy and unequal carryover effects because $\beta_{ij} \neq \beta_{ji}$ and $\lambda_i \neq \lambda$, which thus results in N different solutions from each $S_i(t)$ and requires media-specific costate variables in the optimal control. In Appendix C, we solve the general control problem induced by Equations 19–21 and obtain the optimal N-media IMC strategy:

$$\vec{u}^* = -[M(\vec{\mu})]^{-1}\vec{\mu} \circ \vec{\beta}, \tag{22}$$

where the matrix $M(\vec{\mu})$ is defined in Appendix C, the elements in vector $\vec{\mu}$ are costate variables $\mu_i = m/(1 - \lambda_i + \rho)$, the vector $\vec{\beta} = (\beta_1, ..., \beta_N)'$ contains the media effectiveness parameters, and the symbol $\circ$ denotes element-by-element multiplication of two vectors.

Equation 22 is a closed-form expression for the optimal effort to expend on each medium. It helps managers determine the optimal allocation across N-media in the presence of differential carryover and asymmetric synergy effects. In addition, it helps researchers understand the normative effects of synergy, which we illustrate with two theoretical scenarios.[14]

The first scenario is, How does the two-media allocation rule generalize to the three-media case? To understand this, we evaluate Equation 22 with N = 3 media, obtain the explicit expressions similar to Equations 15 and 16, and determine the allocation ratios $\Lambda_{13} = u_1^*/u_3^*$, $\Lambda_{12} = u_1^*/u_2^*$, and $\Lambda_{23} = u_2^*/u_3^*$. Then, to generalize from the two- to three-media case, ceteris paribus, we let $\beta_{ij} = \kappa$ and $\lambda_i = \lambda$, and we apply comparative dynamics analyses to find that

$$\text{sign}\left(\frac{\partial\Lambda_{ij}}{\partial\kappa}\right) = \begin{cases} - \text{ if } \beta_i > \beta_j, \\ + \text{ if } \beta_i < \beta_j, \text{ and} \\ 0 \text{ if } \beta_i = \beta_j, \end{cases} \tag{23}$$

for all i, j = 1, 2, and 3. Equation 23 indicates that for the three-media case, as synergy increases, advertisers should decrease expenditures on the more effective activity and increase expenditures on the less effective one. In addition,

[13] We thank a reviewer for suggesting this perspective.

[14] We thank a reviewer for suggesting both scenarios.

budget reallocation between equally effective activities is unnecessary even if synergy increases.

In the second scenario, we investigate how managers should change the budget allocation if they experience a large direct effect but no carryover effect for one medium and a small direct effect but a large carryover effect for the other medium. To understand allocation decisions for this setting, we again evaluate Equation 22, this time with $\lambda_1 \neq \lambda_2$. We obtain a new expression for the allocation ratio $\Lambda = u_1^*/u_2^*$, and we conduct comparative dynamics analyses to find that

$$\text{sign}\left(\frac{\partial \Lambda}{\partial \kappa}\right)\Bigg|_{\substack{\beta_1 \to \infty \\ \lambda_1 \to 0 \\ \beta_2 \to 0 \\ \lambda_2 \to 1}} < 0. \tag{24}$$

Equation 24 indicates that managers should decrease the proportion of budget allocated to the more effective medium (i.e., the one with a greater instantaneous effect, $\beta_1 > \beta_2$), thus allocating more than fair share to the less effective medium.

In summary, both scenarios reinforce the fundamental point that budget allocation in the IMC context substantively differs from that based on standard advertising theory. Specifically, the role of synergy (indeed, the quintessential aspect of IMC) is to favor the less effective activities in lieu of the standard allocation principle that advocates budget allocation in proportion to relative effectiveness (see P_1).

DISCUSSION

This section discusses model refinements and suggestions for additional research.[15]

Competition

We investigated how monopolist advertisers (e.g., Intel's Pentium) should allocate media budget strategically by considering cross-media synergy. A natural extension is to determine how budgeting and allocation would change if a competitor's brand advertises (e.g., AMD's Athalon). Such an extension requires differential game theory to solve for optimal strategies (for details, see Erickson 1991).

Uncertainty

In dynamic systems, there are two kinds of uncertainties: observation noise and transition noise. Managers can reduce observation noise by improving their measurement system (e.g., increasing sample size or sampling frequency). Recent research provides methods to mitigate the adverse impact of measurement errors in awareness tracking studies through wavelet filtering (Cai, Naik, and Tsai 2000; Naik and Tsai 2000a). In contrast, transition noise arises partly because of nonconstancy of model parameters and partly because of environmental uncertainty that results from a large number of small events that influence a brand's sales evolution. In our subsequent discussion of dynamic stochastic process functions, we show how managers can determine whether model parameters themselves change systematically or gradually (i.e., slow stochastic variation) over time. Managers can devise optimal marketing strategies in the presence of environmental uncertainty by applying stochastic control theory (for details, see Jagpal 1999; Mantrala, Raman, and Desiraju 1997; Nguyen 1997; Raman 1990; Raman and Chatterjee 1995).

Temporal Aggregation

When data are observed over a time interval (e.g., annual) that is different from the interval implicit in model specification (e.g., monthly), the resulting estimation biases and inefficiency are attributed to temporal aggregation. Hanssens, Parsons, and Schultz (1998, p. 222) and Leeflang and colleagues (2000, p. 279) discuss the marketing literature on this topic. A recent econometric study addresses the problem of random aggregation, in which the data interval itself is uncertain (Jorda 1999). Because this literature predominantly considers linear models, the consequences of temporal aggregation for estimating synergy are not known and require further investigation.

Dynamic Stochastic Process Functions

Remark 2 explains the distinction between response and process functions that Gatignon and Hanssens (1987) introduce. Typically, sales response functions are dynamic, but process functions are not (see Gatignon 1993, p. 703). The estimation and inference of dynamic and stochastic process functions can be carried out as follows: Specifically, brand sales are influenced by communication activities $\vec{u} = (u_1, ..., u_N)'$ with effectiveness parameters $\vec{\beta} = (\beta_1, ..., \beta_N)'$ and synergy parameters $\vec{\kappa} = (\beta_{12}, ..., \beta_{ij}, ..., \beta_{N,N-1})'$. Other variables, such as price, can be introduced by means of a covariate vector $\vec{x}_t$; let $\vec{\gamma}$ contain the direct effects of covariates on sales.[16] All the parameters are stacked into the vector $\vec{\phi} = (\vec{\beta}', \vec{\kappa}', \vec{\gamma}')'$, and each may be affected by some activities in the communications vector $\vec{u}$. Then, the dynamic sales response function is

$$S_t = f(S_{t-1}, \vec{u}_t, \vec{x}_t, \vec{\phi}_t) + \omega_{1t}, \tag{25}$$

and the dynamic stochastic process function is

$$\vec{\phi}_t = p(S_{t-1}, \vec{u}_t, \vec{\phi}_{t-1}) + \omega_{2t}. \tag{26}$$

The lagged vector $\vec{\phi}_{t-1}$ induces process dynamics, and ω_{2t} introduces uncertainty in the evolution of $\vec{\phi}_t$. Such stochastic and time-varying coefficients can arise in models of turbulent markets or new product introductions. We parameterize $f(\cdot)$ and $p(\cdot)$ by incorporating behavioral assumptions; for example, in the work of Naik, Mantrala, and Sawyer (1998, Figures 4 and 9), advertising effectiveness waxes and wanes as a result of the timing and intensity of advertising spending decisions.

The preceding framework permits a variety of dynamic models and error structures. For example, suppose that the error ω_{1t} follows an AR(2) process, $\omega_{1,t} = \rho_1\omega_{1,t-1} + \rho_2\omega_{1,t-2} + \varepsilon_t$. We augment Equation 26 to redefine a new $\tilde{\phi}_t = (\vec{\phi}_t, \omega_{1,t}, \omega_{1,t-1})'$whose transition now includes the two additional rows as follows:

$$\begin{bmatrix} \omega_{1,t} \\ \omega_{1,t-1} \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \omega_{1,t-1} \\ \omega_{1,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}. \tag{27}$$

[15]We thank the reviewers for suggesting these issues.

[16]We acknowledge that the incorporation of price in dynamic sales models is an important issue; this framework can be applied to accomplish that goal.

This example illustrates the flexibility of the framework, which can be applied to estimate models with complex lag structure using any high-order autoregressive integrated moving average (p, d, q) process in general (for details, see Shumway and Stoffer 2000, Ch. 4).

The joint estimation of the response and process models (Equations 25 and 26) is accomplished as follows: We stack sales and the elements in $\vec{\phi}_t$ to form the vector $\tilde{\phi}_t = (S_t, \vec{\phi}_t')'$. Two problems arise because $\tilde{\phi}_t$ is a function of random and dynamic variables. First, the usual maximum-likelihood methods are not applicable because they assume parameters to be fixed (i.e., nonrandom constants). Second, other maximum-likelihood approaches (e.g., random coefficients model) do not apply because they require that distributions of parameters are stationary (i.e., their moments are non-dynamic constants). Both problems are resolved by filtering theory (see Jazwinski 1970), which permits parameters to evolve as nonstationary, dynamic, stochastic processes. Thus, we can apply the Kalman filter to estimate dynamic models with both time-varying and stochastic parameters with general error structures (e.g., autoregressive integrated moving average) underlying sales data. When the functions $f(\cdot)$ and $p(\cdot)$ are linear in S and ϕ, we apply the standard Kalman filter used in this study; if not, we use the extended Kalman filter (see Shumway and Stoffer 2000).

Overparameterization

When a brand uses several communications activities, the number of parameters in the model increases rapidly, thus resulting in overparameterization. For example, a brand that uses 10 media would require 10 effectiveness and 45 synergy parameters (i.e., two-way interactions), thus requiring large sample sizes for model calibration. Furthermore, media-specific carryover effects are unidentified and thus inestimable unless sales by each medium are known (see Remark 4). In addition, grid search in high-dimensional space is impractical. Moreover, even if these pragmatic estimation problems were overcome, the resulting model would be grossly overparameterized, causing it to generalize poorly to new situations. Thus, we recommend keeping carryover effect constant in empirical applications. To reduce overparameterization further, we suggest two approaches: dimension reduction and variable selection.

In dimension reduction, we construct new variables as linear combinations of original variables (including interactions). Factor analysis or principal components analysis achieves dimension reduction but ignores information in the dependent variable while constructing the factors. In contrast, sliced inverse regression finds the factor structure by incorporating information from both the dependent and the independent variables. Naik, Hagerty, and Tsai (2000) discuss the relative efficacy of various dimension-reduction methods.

In variable selection, we attempt to identify a small set of important variables by using information criteria (e.g., AIC_C). When overparameterization becomes severe (i.e., the number of variables exceeds the sample size), most estimation methods break down. In such cases, at least in principle, the partial least squares (PLS) approach enables model calibration (see Martens and Naes 1989, p. 136). Naik and Tsai (2000b) demonstrate that PLS yields consistent estimates even if model specification is unknown or incorrect. They further show that PLS is robust in the presence of strong multicollinearity, which could arise when models include cross-media interactions. Given the desirable properties of PLS, a strategy is to apply PLS to estimate an overparameterized model and then select relevant variables using the information criterion, derived by Naik and Tsai (2001), which correctly penalizes both misspecification and overparameterization. By applying these techniques, further research can elucidate finite-sample properties in simulated and real data sets.

Single-Index Models

A general sales model can be specified as an autoregressive process:

$$S_t = f(S_{t-1}, u_1, u_2, ..., u_N) + \varepsilon_t. \tag{28}$$

However, Equation 28 serves as a conceptual expression: It cannot be estimated empirically or analyzed theoretically without making any additional assumptions on the response function $f(\cdot, ..., \cdot)$ with N + 1 variables. If it is assumed that the response function is continuous and that its (N + 1) variables can be combined linearly, the resulting model is known as the single-index model (for more information, see Horowitz 1998). Naik and Tsai (2001) have developed an approach for the estimation and selection of single-index models. The linear index combining all the variables in Equation 28 lends an appealing interpretation: It represents the firm's total communication effort across N-media and captures the past effort through the lagged sales term. We emphasize that besides the nonlinear function f, the linear index itself can include various nonlinear effects by means of interactions and/or power transformations of variables. Finally, because single-index models are not yet applied in marketing (Leeflang et al. 2000, p. 403), they present new frameworks to advance empirical and theoretical understanding of multi- and cross-media communications.

CONCLUSION

The IMC framework demands a new perspective of the measurement and optimization of a firm's promotional mix (Schultz 2002). A central tenet of the IMC approach, which distinguishes it from the conventional view, is that each medium enhances the contributions of all other media. This distinction is driven by the potential existence of synergy, that is, the added value of one medium as a result of the presence of another medium, causing the combined effect of media to exceed the sum of their individual effects. Thus, the combined impact of multimedia activities such as television, print, radio, Internet, direct response, sales promotion, and public relations can be much greater than the sum total of their individual effects.

Although many marketers and advertising agencies embrace this concept, perhaps because of its intuitive appeal, research on the role of synergy in multimedia communications is scarce. Therefore, we present empirical and theoretical analyses of the effects of synergy in the IMC context. Both analyses augment the advertising literature. Specifically, we furnish empirical evidence on the existence of synergy between television and print advertising in consumer markets. Our estimation methodology can be applied to brand-specific data to estimate and infer the effectiveness of multimedia communications and synergy among them. We derive several propositions to elucidate the effects of synergy on media budget, media mix, and advertising carryover. We show that as synergy increases, the advertiser requires a larger media budget. Furthermore, a smaller (larger) proportion of this media budget should be allocated to the more (less) effective activity. These findings have important implications for advertising overspending. Specifically, a normative model that ignores synergy understates the optimal budget, and thus the comparison of actual with optimal spending levels exaggerates the magnitude of overspending.

In conclusion, we emphasize that managers can use market data to estimate and infer not only media effectiveness but also cross-media synergy. We hope that the proposed models and methods assist managers in integrating their multimedia campaigns.

APPENDIX A: DERIVATION OF THE OPTIMAL IMC STRATEGY

Our objective is to maximize Equation 13:

$$J(u, v) = \int_0^\infty e^{-\rho t}\Pi[S(t), u(t), v(t)]dt,$$

where $\Pi(S, u, v) = mS - u^2 - v^2$, subject to Equation 14:

$$\frac{dS}{dt} = \beta_1 u(t) + \beta_2 v(t) + \kappa u(t)v(t) - (1 - \lambda)S(t).$$

We define the Hamiltonian $H(u, v, \mu)$:

$$\text{(A1)} \qquad H(u,v,\mu) = (mS - u^2 - v^2) + \mu\,[\beta_1 u + \beta_2 v + \kappa uv - (1 - \lambda)S],$$

where μ is the costate variable. At optimality, the necessary conditions are

$$\text{(A2)} \qquad \frac{\partial H}{\partial u} = 0, \quad \frac{\partial H}{\partial v} = 0, \text{ and} \quad \frac{d\mu}{dt} = \rho\mu - \frac{\partial H}{\partial S}.$$

Furthermore, these necessary conditions are sufficient because the Hamiltonian $H(\cdot)$ is concave in u and v.

We determine u and v in terms of the costate variable μ from the first two equations in Equation A2. Thus, we get

$$\text{(A3)} \qquad \frac{\partial H}{\partial u} = 0 \Rightarrow -2u + \beta_1\mu + \kappa\mu v = 0, \text{ and} \quad \frac{\partial H}{\partial v} = 0 \Rightarrow -2v + \beta_2\mu + \kappa\mu u = 0.$$

Applying Cramer's Rule, we express the controls u and v in terms of the costate variable μ:

$$\text{(A4)} \qquad u^* = \frac{\mu(2\beta_1 + \mu\beta_2\kappa)}{4 - \mu^2\kappa^2}, \text{ and} \quad v^* = \frac{\mu(2\beta_2 + \mu\beta_1\kappa)}{4 - \mu^2\kappa^2}.$$

We then express μ in terms of model parameters by using the costate equation and transversality conditions. The costate equation is given by

$$\text{(A5)} \qquad \frac{d\mu}{dt} = \rho\mu - \frac{\partial H}{\partial S} \Rightarrow \frac{d\mu}{dt} = -m + \mu(1 - \lambda) + \rho\mu.$$

For an autonomous system with infinite time horizon, the transversality conditions are obtained from the steady-state conditions on the state and costate variables (Kamien and Schwartz 1991, p. 160), which are given by

$$\text{(A6)} \qquad \frac{\partial S}{\partial t} = 0, \text{ and} \quad \frac{\partial \mu}{\partial t} = 0.$$

Solving Equations A5 and A6 simultaneously, we obtain

$$\text{(A7)} \qquad \mu(t) = \frac{m}{(1 - \lambda + \rho)}.$$

Finally, we substitute Equation A7 in Equation A4 to find the optimal controls

$$\text{(A8)} \qquad u^* = \frac{m[\beta_2\kappa m + 2\beta_1(1 + \rho - \lambda)]}{4(1 + \rho - \lambda)^2 - \kappa^2 m^2}, \text{ and} \quad v^* = \frac{m[\beta_1\kappa m + 2\beta_2(1 + \rho - \lambda)]}{4(1 + \rho - \lambda)^2 - \kappa^2 m^2}.$$

Because the optimal controls take positive values, we have

$$\text{(A9)} \qquad 4 - \left(\frac{m\kappa}{1 + \rho - \lambda}\right)^2 > 0.$$

Thus, the optimal IMC strategy is specified by Equations A8 and A9.

APPENDIX B: PROOFS OF THE PROPOSITIONS

P_1

Let $\kappa = 0$. We differentiate Equation 15 with respect to β_1 and find that the sign $[\partial u^*/\partial\beta_1] = \text{sign}\{4 - [\kappa m/(1 + \rho - \lambda)]^2\} > 0$. The last inequality follows from Equation A9. Similarly, we differentiate Equation 16 with respect to β_2 and use Equation A9 to find that sign $[\partial v^*/\partial\beta_2] > 0$. Next, differentiating Equation 17 with respect to β_i and using Equation A9, we show that sign $[\partial\beta/\partial\beta_i] > 0$ for i = 1, 2. Finally, substituting $\kappa = 0$ in Equation 18, we note that $\Lambda = \beta_1/\beta_2$ for any values of λ, ρ, and m.

P_2

Differentiating Equation 17 with respect to κ, we obtain

$$\text{(B1)} \qquad \frac{\partial B}{\partial \kappa} = \frac{m^2(\beta_1 + \beta_2)}{[2(1 + \rho - \lambda) - m\kappa]^2} > 0.$$

P_3

We differentiate Equation 18 and observe that sign $[\partial\Lambda/\partial\kappa] = \text{sign}\,[\beta^2{}_2 - \beta^2{}_1]$, which is negative when $\beta_1 > \beta_2$ and positive when $\beta_1 < \beta_2$. When $\beta_1 = \beta_2$, $\Lambda = 1$ regardless of the magnitude of κ.

P_4

Differentiating Equation 17 repeatedly, we find that sign $[\partial B/\partial\lambda] > 0$ and sign $[\partial^2 B/\partial\lambda\partial\kappa] > 0$. When $\kappa = 0$, Equation 18 reveals that $\Lambda = \beta_1/\beta_2$, which is not a function of λ. When $\kappa \neq 0$, the allocation ratio in Equation 18 functionally depends on λ. In this case, sign $[\partial\Lambda/\partial\lambda]$ = sign $[\beta^2{}_2 - \beta^2{}_1]$, which is negative when $\beta_1 > \beta_2$ and positive when $\beta_1 < \beta_2$.

APPENDIX C: DERIVATION FOR N-MEDIA IMC STRATEGY

Letting $S(t) = \Sigma_{i=1}^{i=N} S_i(t)$, we define the Hamiltonian $H(\vec{u}, \vec{\mu}, S)$:

$$H(\vec{u}, \vec{\mu}, S) = \left(mS - \sum_{i=1}^{N} u_i^2\right) + \sum_{i=1}^{N} \mu_i \left[\beta_i u_i + \sum_{\substack{i=1\\ i\neq j}}^{N} \beta_{ij} u_i u_j - (1-\lambda_i) S_i\right], \tag{C1}$$

where the costate variable μ_i is in $\vec{\mu} = (\mu_1, ..., \mu_N)'$. At optimality, we obtain 2N necessary conditions:

$$\frac{\partial H}{\partial \vec{u}} = 0, \quad \text{and} \quad \frac{d\mu_i}{dt} = \rho\mu_i - \frac{\partial H}{\partial S_i}. \tag{C2}$$

The equation $\partial H/\partial\vec{\mu} = 0$ generates N first-order conditions for the optimal levels of each medium in $\vec{\mu}$:

$$\begin{aligned} \frac{\partial H}{\partial u_1} &= 0 \Rightarrow -2u_1 + \beta_{12}\mu_1 u_2 + ... + \beta_{1N}\mu_1 u_N = -\mu_1\beta_1 \\ \frac{\partial H}{\partial u_2} &= 0 \Rightarrow \beta_{21}\mu_2 u_1 - 2u_2 + ... + \beta_{2N}\mu_2 u_N = -\mu_2\beta_2 \\ &\vdots \qquad \vdots \qquad\qquad\qquad \vdots \\ \frac{\partial H}{\partial u_N} &= 0 \Rightarrow \beta_{N1}\mu_N u_1 + \beta_{N2}\mu_N u_2 + ... -2u_N = -\mu_N\beta_N. \end{aligned} \tag{C3}$$

We define an $N \times N$ matrix M that depends on $\vec{\mu}$ as

$$M(\vec{\mu}) = \begin{pmatrix} -2 & \beta_{12}\mu_1 & \cdots & \beta_{1N}\mu_1 \\ \beta_{21}\mu_2 & -2 & \cdots & \beta_{2N}\mu_2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{N1}\mu_N & \beta_{N2}\mu_N & \cdots & -2 \end{pmatrix}, \tag{C4}$$

and we denote the media effectiveness vector by $\vec{\beta} = (\beta_1, ..., \beta_N)'$ so that Equation C3 becomes the following:

$$M(\vec{\mu})\vec{u} = -\vec{\mu} \circ \vec{\beta}, \tag{C5}$$

where the symbol $\circ$ denotes element-by-element multiplication of two vectors. Thus, provided that $M(\vec{\mu})$ is nonsingular, the optimal strategy is given by

$$\vec{u}^* = -[M(\vec{\mu})]^{-1}\vec{\mu} \circ \vec{\beta}. \tag{C6}$$

Finally, solving $d\mu_i/dt = \rho\mu_i - (\partial H/\partial S_i)$, we obtain

$$\mu_i(t) = \frac{m}{(1 - \lambda_i + \rho)}, \tag{C7}$$

which when substituted in Equation C6 provides the optimal N-media IMC strategy as a function of model parameters.

REFERENCES

Belch, George E. and Michael A. Belch (1998), *Advertising and Promotion: An Integrated Marketing Communications Perspective*, 4th ed. Boston: Irwin-McGraw Hill.

Bucklin, Randolph E. and Sunil Gupta (1999), "Commercial Use of UPC Scanner Data: Industry and Academic Perspectives," *Marketing Science*, 18 (3), 247–73.

Cai, Zongwu, Prasad A. Naik, and Chih-Ling Tsai (2000), "Denoised Least Squares Estimators: An Application to Estimating Advertising Effectiveness," *Statistica Sinica*, 10 (4), 1231–41.

Dorfman, Robert and Peter O. Steiner (1954), "Optimal Advertising and Optimal Quality," *American Economic Review*, 44 (December), 826–36.

Edell, Julie E. and Kevin L. Keller (1999), "Analyzing Media Interactions: The Effects of Coordinated TV-Print Advertising Campaigns," Working Paper No. 99-120, Marketing Science Institute, Cambridge, MA.

Engle, R.F., D.F. Hendry, and J.-F. Richard (1983), "Exogeneity," *Econometrica*, 51 (2), 277–304.

Enrico, Dottie (1996), "Dockers' Pants Ads Don't Appear to Have Legs," *USA Today*, (July 22).

Erickson, Gary M. (1991), *Dynamic Models of Advertising Competition: Open- and Closed-Loop Extensions.* Boston: Kluwer Academic Publishers.

Feichtinger, Gustav, Richard F. Hartl, and Suresh P. Sethi (1994), "Dynamic Optimal Control Models in Advertising: Recent Developments," *Management Science*, 40 (2), 195–226.

Gatignon, Hubert (1993), "Marketing Mix Models," in *Handbook of OR and MS*, Vol. 5, J. Eliashberg and G.L. Lilien, eds. New York: North-Holland, 697–732.

——— and Dominique M. Hanssens (1987), "Modeling Marketing Interactions with Application to Salesforce Effectiveness," *Journal of Marketing Research*, 24 (August), 247–57.

Gopalakrishna, Srinath and Rabikar Chatterjee (1992), "A Communications Response Model for a Mature Industrial Product: Application and Implications," *Journal of Marketing Research*, 29 (May), 189–200.

Hamilton, James D. (1994), *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Hanssens, Dominique M., Leonard J. Parsons, and Randall L. Schultz (1998), *Market Response Models: Econometric and Time Series Analysis*. Boston: Kluwer Academic Publishers.

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

Horowitz, Joel L. (1998), *Semiparametric Methods in Econometrics*. New York: Springer.

Hurvich, Clifford M. and Chih-Ling Tsai (1989), "Regression and Time Series Model Selection in Small Samples," *Biometrika*, 76 (2), 297–307.

Jagpal, Harsharanjeet (1981), "Measuring Joint Advertising Effects in Multiproduct Firms," *Journal of Advertising Research*, 21 (1), 65–69.

Jagpal, Sharan (1999), *Marketing Strategy and Uncertainty*. New York: Oxford University Press.

Jazwinski, Andrew H. (1970), *Stochastic Processes and Filtering Theory*. San Diego: Academic Press.

Jorda, Oscar (1999), "Random-Time Aggregation in Partial Adjustment Models," *Journal of Business & Economic Statistics*, 7 (3), 382–95.

Kamien, Morton and Nancy Schwartz (1991), *Dynamic Optimization*. New York: Elsevier Science Publishing Company.

Keller, Kevin L. (1998), *Strategic Brand Management*. Upper Saddle River, NJ: Prentice Hall.

Kuehn, Alfred A. (1962), "How Advertising Performance Depends on Other Marketing Factors," *Journal of Advertising Research*, 2 (March), 2–10.

Leeflang, Peter, Dick R. Wittink, Michel Wedel, and Phillip A. Naert (2000), *Building Models for Marketing Decisions*. Boston: Kluwer Academic Publishers.

Little, John D.C. (1979), "Aggregate Advertising Models: The State of the Art," *Operations Research*, 27 (4), 629–67.

Mantrala, Murali K. (2002), "Allocating Marketing Resources," in *Handbook of Marketing*, Barton A. Weitz and Robin Wensley, eds. Thousand Oaks, CA: Sage Publications.

———, Kalyan Raman, and Ramarao Desiraju (1997), "Sales Quota Plans: Mechanisms for Adaptive Learning," *Marketing Letters*, 8 (4), 393–405.

Martens, Harald and Tormod Naes (1989), *Multivariate Calibration*. New York: John Wiley & Sons.

Montgomery, David B. and Alvin Silk (1972), "Estimating Dynamic Effects of Marketing Communications Expenditures," *Management Science*, 18 (10), B485–B501.

Naik, Prasad A., Michael R. Hagerty, and Chih-Ling Tsai (2000), "A New Dimension Reduction Approach for Data-Rich Marketing Environments: Sliced Inverse Regression," *Journal of Marketing Research*, 37 (February), 88–101.

———, Murali K. Mantrala, and Alan Sawyer (1998), "Planning Pulsing Media Schedules in the Presence of Dynamic Advertising Quality," *Marketing Science*, 17 (3), 214–35.

——— and Chih-Ling Tsai (2000a), "Controlling Measurement Errors in Models of Advertising Competition," *Journal of Marketing Research*, 37 (February), 113–24.

——— and ——— (2000b), "Partial Least Squares Estimator for Single-Index Models," *Journal of the Royal Statistical Society*, Series B, 62 (4), 763–71.

——— and ——— (2001), "Single-Index Model Selections," *Biometrika*, 88 (3), 821–32.

Nerlove, Marc and Kenneth Arrow (1962), "Optimal Advertising Policy Under Dynamic Conditions," *Economica*, 29 (May), 129–42.

The New York Times (1994), "DDB Needham Copenhagen Takes Top Honors," (May 10), C18(N), D22(L).

Nguyen, Dung (1997), *Marketing Decisions Under Uncertainty*. Boston: Kluwer Academic Publishers.

Oud, J.H.L., R.A.R.G. Jansen, and D.M.A. Haughton (1999), "Small Samples in Structural Equation State Space Modeling," in *Statistical Strategies for Small Sample Research*, R. Hoyle, ed. Thousand Oaks, CA: Sage Publications, 288–305.

Palda, Kristian S. (1964), *The Measurement of Cumulative Advertising Effects*. Englewood Cliffs, NJ: Prentice Hall.

Parsons, Leonard J. (1974), "An Econometric Analysis of Advertising, Retail Availability, and Sales of a New Brand," *Management Science*, 20 (February), 938–47.

Petersen, Laurie (1991), "Northwestern Launches 'Integrated' 101," *Adweek's Marketing Week*, 32 (39), 7.

Raman, Kalyan (1990), "Stochastically Optimal Advertising Policies Under Dynamic Conditions: The Ratio Rule," *Optimal Control Applications and Methods*, 11 (3), 283–88.

——— and Rabikar Chatterjee (1995), "Stochastic Optimal Control of Monopoly Pricing Models with Dynamic Demand and Production Cost," *Management Science*, 41 (1), 144–62.

Rao, C.R. (1973), *Linear Statistical Inference and Its Applications*, 2d ed. New York: John Wiley & Sons.

Rothenberg, Thomas J. (1971), "Identification in Parametric Models," *Econometrica*, 39 (3), 577–91.

Schultz, Don E. (1993), "Integrated Marketing Communications: Maybe Definition Is in the Point of View," *Marketing News*, 27 (2), 17.

——— (2002), "Outdated Approaches to Planning Needs Revamping," *Marketing News*, 36 (23), 6–7.

Sethi, Suresh P. (1977), "Dynamic Optimal Control Models in Advertising: A Survey," *SIAM Review*, 19 (4), 685–725.

Shumway, Robert H. and David S. Stoffer (2000), *Time Series Analysis and Its Applications*. New York: Springer-Verlag.

Swinyard, William R. and Michael L. Ray (1977), "Advertising–Selling Interactions: An Attribution Theory Experiment," *Journal of Marketing Research*, 14 (November), 509–516.

Zellner, Arnold and Martin S. Geisel (1970), "Analysis of Distributed Lag Models with Applications to Consumption Function Estimation," *Econometrica*, 38 (6), 865–88.