Naik, Prasad A.; Peters, Kay; Raman, Kalyan

Working Paper

Heterogeneous Response Functions in Advertising

Arbeitspapiere des Lehrstuhls für Innovation, Neue Medien und Marketing

Provided in cooperation with:
Christian-Albrechts-Universität Kiel (CAU)

Heterogeneous Response Functions in Advertising

Prasad A. Naik
Professor of Marketing and Chancellor's Fellow
Graduate School of Management
University of California Davis
One Shields Avenue, AOB IV, Room 151
Davis, CA 95616
Phone: (530) 754 9834
Fax: (530) 752 2924
Email: panaik@ucdavis.edu

Kay Peters
Managing Director
Center for Interactive Marketing and Media Management
c/o Institute of Marketing, University of Muenster
Am Stadtgraben 13-15, 48143 Muenster, Germany
Phone: +49 (251) 83-25025
Fax: +49 (251) 83-25024
Email: kay.peters@uni-muenster.de

Kalyan Raman
Professor IMC, Medill School
Courtesy Professor of Marketing, Kellogg School of Management,
Northwestern University
Evanston, IL 60208
Phone: (847) 467 0847
Fax: (847) 491 5925
k-raman@northwestern.edu

October 2008

Authors are listed in alphabetical order. Prasad A. Naik acknowledges the financial support received from the Chancellor's Fellowship at the University of California at Davis. Kay Peters acknowledges the financial and data support of the Siegfried Vögele Institute, Bonn and access to data from Research International, Hamburg. We benefited from helpful discussions at the Marketing Dynamics Conference 2007 in Groningen.

Heterogeneous Response Functions in Advertising

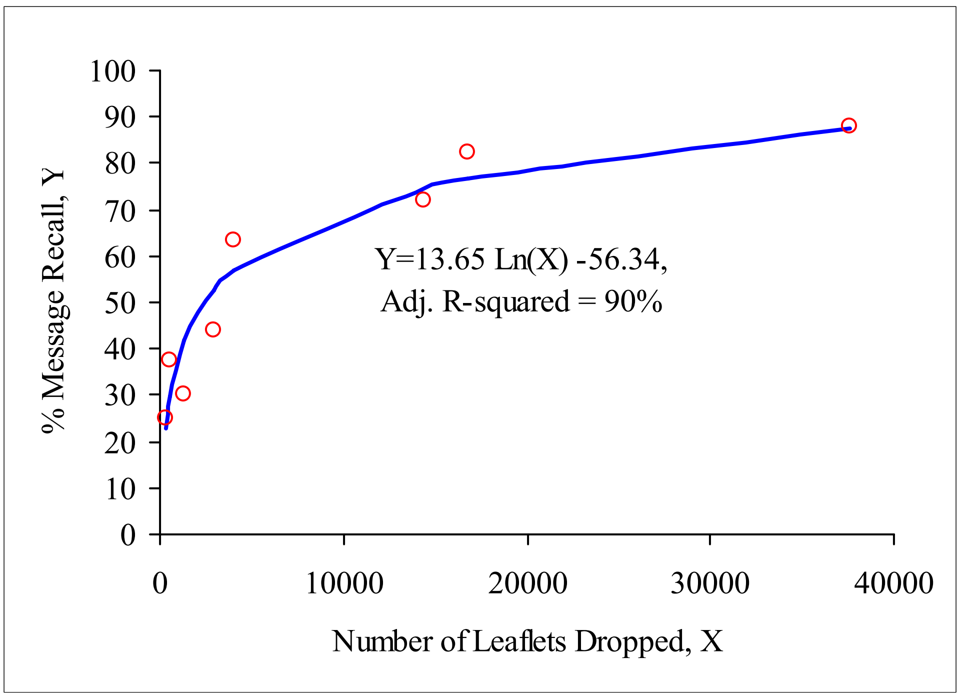
Abstract:

De Fleur (1956) provides the earliest evidence of diminishing returns. He finds a common logarithmic pattern for leaflets dropped and message recalled in field experiment. Since then, many researchers have applied logarithmic or square root patterns to capture the effect of diminishing returns with their advertising response modeling across different media. But discussions with managers support the notion that the diminishing returns to incremental dollars spent on one medium (say, television) are not likely to be the same as those for equivalent dollars spent on other media (e.g., Print). But if diminishing returns indeed vary across media, how does that change the resulting allocation recommendation? To address this issue, we derive a dynamic model that captures the notion of differential diminishing returns and disentangles it from closely related notions of differential carryovers and differential ad effectiveness. Second, we develop a systematic method to estimate the model's parameters using market data and illustrate empirically that all three effects, diminishing returns, carryover and ad effectiveness vary across the four media employed. Finally, we investigate the normative implications for managerial decision-making. Here, we additionally account for varying media buying efficiencies across media. Taken together, the approach and its illustration should provide managers with a better toolkit to allocate their multimedia budgets.

1. Introduction

De Fleur (1956) provides the earliest evidence of diminishing returns. He painstakingly collects field experimental data with the U.S. Air Force on the number of leaflets dropped and the message recalled across eight similar towns. Relating two types of response across these towns to Weber's law - originally developed for individual stimulus response -, he finds a common logarithmic pattern for these communities (cf. Figure 1). Since then, many researchers have applied logarithmic or square root patterns to capture the effect of diminishing returns with their advertising response modeling. Most of these studies continue to infer superior budget allocations from their applications. For the sake of model parsimony, diminishing returns are often assumed to follow an equal shape across media or measured jointly with advertising effectiveness through elasticities.

Figure 1. De Fleur's (1956) Law of Diminishing Returns



In contrast, although De Fleur (1956, p.13) confirms a logarithmic shape of diminishing returns for his leaflet study, he suggests that other media may exhibit different shapes due to varying characteristics across media. Also, discussions with marketing managers support the notion that the diminishing returns to incremental dollars spent on one medium (say, television) are not likely to be the same as those for equivalent dollars spent on other media (e.g., Print or Billboards). But if diminishing returns indeed vary across media, how does that change the resulting allocation recommendation? However, there exists no marketing study that either empirically validates this claim or provides a systematic approach for managers to verify this possibility. A reason for this absence is the complexity that results from three confounding sources of media spending impact:

(i) carryovers varying by media

(ii) ad effectiveness varying by media, and

(iii) potentially different degrees of diminishing returns across media.

To separate these phenomena, we need to develop a dynamic model that captures the notion of differential diminishing returns and disentangles it from closely related notions of differential carryovers and differential ad effectiveness. This model development constitutes the primary contribution of this article. The second contribution is to develop a systematic method to estimate the model's parameters using market data and then to illustrate empirically that the diminishing returns do differ across different class of media such as Broadcast (i.e., TV and Radio), Print (i.e., Newspapers and Magazines), and Below-the-Line (e.g., Cinema, Billboards). Finally, the third contribution is to investigate the normative implications for managerial decision-making. Here, we additionally account for varying media buying efficiencies across media. For example, how should managers change the budget allocation as the degree of diminishing returns increases, holding the effects of ad effectiveness, carryover, and media buying efficiency constant.

Applying our new approach to the softdrink market, we demonstrate the ability of our model and estimation approach to disentangle the various effects of advertising in a multimedia environment. We find that all three effects, diminishing returns, carryover and ad effectiveness as well as media buying efficiencies indeed vary across the four media employed. Our Maximum Likelihood estimation of a Kalman Filter provides robust inferences and yields satisfactory cross validation and one-period-ahead forecast results. Using the Hamiltonian and comparative static analysis we derive profit maximizing allocation rules that yield interesting allocation recommendations. Taken together, the approach and its illustration should provide managers with a better toolkit to allocate their multimedia budgets.

The balance of the paper continues with a literature review on multimedia communication modeling in §2. We derive the model and the allocation problem in §3. In §4, we describe the data of our application and develop an appropriate estimation approach based on the Kalman Filter with robust inferences. The diagnostics and empirical results are discussed in §5, followed by the normative investigation in §6. We conclude with a summary in §7.

2. Literature Review

In a recent review article on advertising Vakratsas (2005) stated that the marginal response to each medium and the shape of such responses may differ across media, but that there is a gap in addressing these important issues in the literature. Additionally, he calls for further research into varying media efficiencies (cf. also MSI 2002) and on non-traditional media (e.g., the internet and billboards). Following up on his statement and illustrating the gap that our contribution addresses, we review the literature for relevant studies covering more than one advertising medium. Accordingly, this approach excludes studies on marketing-mix issues that incorporate advertising as a combined variable across all media. Furthermore, we did not consider detailing or sampling as advertising media (e.g., Chintagunta and Vilcassim 1994; Naik, Raman, and Winer 2005). We also excluded all contributions that did not have an empirical application, e.g. Raman and Naik (2004), and most papers on media planning issues where a simulation provides an illustration to the model developed (e.g., Bass and Lonsdale 1966; Aaker 1975; Little and Lodish 1969; Srinivasan 1976). Overall, that left us with 18 studies that fit our criteria. These studies have been evaluated on application data, media covered, modeling approach and specified media effects. Additionally, we examined them on their estimation approach and extent of a normative analysis. Table 1 reflects our comprehensive review and exhibits the differences across the studies concerning these criteria. Followingly, we integrate our review with a brief summary of related contributions that provide arguments on why these effects may differ among different media.

Differential Diminishing Returns

DeFleur (1956) provides the earliest evidence of diminishing returns by painstakingly collecting field experimental data with the U.S. Air Force on the number of leaflets dropped and the message recalled. As billions of leaflets have been dropped in various wars, he argues that from sheer numbers alone, leaflets can be regarded as a mass medium. Like other media, leaflet operations are generally designed to accomplish a fast spreading of the message trough a population, thus reinforcing, modifying or converting attitudes of the audience. The hypothesis central to his research is that the relationship between stimulus intensity and response would be one of diminishing returns. That is, although dropping more leaflets on a community generates greater response, the increments in response are expected to become smaller as additional units of the stimulus are employed. For the experiment, he carefully selected eight small towns in the state of Washington nearly as alike as possible. Each of the towns would receive another intensity of leaflet dropping, ranging from 1 per four persons to 32 per person. In order to generate comparable conditions in the experiment for all sites, he instructed the U.S. Air Force to take aerial photographs, construct maps and make reconnaisance flights to ensure an even distribution of the leaflets throughout the towns. He also aligned the local mass media to refrain from reporting on the experiment for at least three days after the dropping. After three days, at exactly the same time teams of two interviewers went to several sites within the towns to check on the number of people informed on the content of the the leaflets. Additionally, the leaflets contained separable slips that could be mailed to a specified address. For both response measures, given the varying stimuli across towns, he suggested the diminishing returns relate to the pattern of Weber's law. That has been originally derived by Fechner (1851, 1887) for individual stimulation suggesting logarithmic diminishing returns. Figure 1 shows De Fleur's original data and the fitted "Law of Diminishing Returns." But as DeFleur (1956, p. 13) already states, every medium of commmunication is thought to have certain advantages not shared in all degrees by other media. It is thus likely, that different media may yield different diminishing returns.

According to our review of the relevant literature (cf. Table 1), the studies employing an additive response function consider diminishing returns, if specified, either to follow a logarithmic or square-root pattern across all media (e.g., Doyle and Saunders 1990; Naik, Schultz and Srinivasan 2007). Those five studies that use a multiplicative response function have elasticities representing both effects, diminishing returns and ad effectiveness, simultaneously (e.g., Montgomery and Silk 1972). Dertouzos and Garber (2006) assume an s-shape response function but do not test their assumption empirically. Another recent approach is represented by Dekimpe and Hanssens (1995) on the aggregate and Ansari, Mela, and Neslin (2008) on the individual level. Both studies use VAR models that implicitly assume variable shapes of diminishing returns. With this approach, an additional effort has to be invested to recover the shape of the media specific dimishing returns functions. But it may be difficult to disentangle this effect from the also jointly considered carryover effects respectively. Summarizing these findings from the literature, only Jagpal (1981) specifies and tests differential diminishing returns across two media explicitly, although in a static analysis that neglects the important dynamics of advertising.

Differential Ad Effectiveness

It is interesting to note that in contrast to diminishing returns, ad effectiveness has always been assumed to vary by media across all the studies conveyed. As mentioned above, in five studies it has been represented jointly with diminishing returns effects by a single medium elasticity.

Differential Carryover

Since Scott (1903) published his work on the psychology of advertising, several investigations into the nature and determinants of carryover effects have been commissioned. One explanation derived from early experiments on the human memory processing is that advertising impressions accumulate over time. This effect supposedly works in two ways as repetition enforces the previous impact while fading of the impression over time is simultaneously reduced (e.g., Lucas and Britt 1950). Another explanation is based on the interference of advertising messages with a given mindset, thus inducing an update of processed cumulative knowledge or a partial adjustment.

Several early studies measure carryovers of advertising explicitly (e.g., Nerlove and Waugh 1961; Palda 1965; Tull 1965). Most studies investigated a single carryover effect for the combined advertising spend (compare Leone 1995 for an overview), although other early contributions (e.g., Gensch 1970; Jastram 1955) and practitioners indicate that carryover might indeed vary with by advertising medium. Berkowitz, Allaway, and D'Souza (2001b) summarize the literature on why differential lags between media may exist. One stream of argument relates to the varying cognitive impact of different media, such as the rate of forgetting being inversely related to the degree the message has been learnt (e.g., Bean 1912). For alternative media with different characteristics learning exhibits different intensities, and accordingly also the rate of forgetting may vary by media. A second line of argument is based on psychological studies of human memory processing, which depends on the holistic presentation of the (advertising) message with special attention to the sensory stimuli conveyed. As those stimuli involved in processing the message vary across media, so may the rate of information take-up and forgetting. Extending some previous research into differential carryovers (e.g., Montgomery and Silk 1972), Berkowitz, Allaway, and D'Souza (2001b) continue to evaluate the profit impact of (mis-)specifying only a single carryover across various media. Using a simulation study, they show that as differential lags between media increase, also the optimal media budget allocation changes. Specifying a single carryover across all media thus leads to suboptimal allocation recommendations. Accordingly, we need to test whether carryovers indeed vary with media.

Of the eighteen studies reviewed here, just five applications do not specify a carryover at all. The remaining studies consider at least a common carryover across all media. But only six studies specify and confirm the existence of differential carryovers across media (e.g., Montgomery and Silk 1972; Hanssens and Levien 1983). Differential carryovers have been implicitly specified within the VAR frameworks of Dekimpe and Hanssens (1995), Vakratsas and Ma (2005) as well as within the Baysian approach by Ansari, Mela, and Neslin (2008). Here, the medium-specific long-term impacts have been derived via impulse response functions holding other effects constant. While having to order the sequence of impacts, it is thus feasible to derive joint differential impacts with some additional effort.

Summarizing our review across all three medium-specific effects, we have to contemplate that in each of the studies one or another effect has not been accounted for. In other words, there is no study or approach that jointly investigates or accounts for differential ad effectiveness, differential carryovers and differential diminishing returns. Thus, to varying degrees there is some confounding of these effects in all reviewed multimedia studies. Accordingly, our approach addresses a gap in the literature that represents an increasingly important challenge for advertising managers in the current multimedia environment. We will continue with developing a model that disentangles these three media-specific effects in the following section.

Table 1. Main Differences Among Related Studies

Author(s) (Year)	Data			Model			Media Effects					Estimation		Normative Analysis
				Response Function	Analysis	Level	Dim. Return	Ad effectiveness	Carryover		Confounding of effects			
	Industry (Detail) Period (Time frame) (R)egions (L)ocal	# different media	unit of media	(M)ultiplicative (A)dditive (L)ogistic	(S)tatic (D)ynamic	(A)ggregate (I)nidividual	Shape (Specification) (A)ssumed/(T)ested	Yes / No	Yes / No	by media?		Method Cross Validation Forecasting	Robust Inference	Yes/No Comp. Stat.
Montgomery/Silk (1972)	Ethical Drug (n.a.) 54m (n.a.)	3	USD	M	D	A	Decreasing Elasticity (T)	with Elasticity	Yes	Yes	Yes	OLS No No	No	No No
Hugues (1975)	Car (Renault5) n.a. (n.a.) France	3	USD / opp. of contact	n.a.	S	A	No	Yes	No		Yes	Monte Carlo No No	No	No No
Jagpal (1981)	Banking (n.a.) 33m (1976-1978)	2	USD	A	S	A	Decreasing Elasticity (T)	Yes	No		Yes	PLS No No	No	No No
Hanssens/Levien (1983)	Military (n.a.) 36m (1976-78) 43R	6	USD pc	M	(D)	A	Decreasing Elasticity (T)	with Elasticity	in 2 models	in 2 models	Yes	OLS Yes No	No	No No
Carroll et al (1985)	Military (n.a.) 1y (1979-80) 26R	5	USD pc	M	D	A	Decreasing Elasticity (T)	with Elasticity	No		Yes	OLS No No	No	No No
Eastlack/Rao (1986)	Softdrink (V8-Cocktail) 3m (1975) 2R 118w (1975-77) 33R	2	GRP pc	A	D	A	No	Yes	Yes	Yes	Yes	OLS No No	No	No No
Gatignon/Hanssens (1987)	Military (n.a.) 36m (1976-78) 6R	2	USD	M	S	A	Decreasing Elasticity (T)	with Elasticity	No		Yes	GLS SUR Yes No	No	(Yes) for result only for 2 dimensions
Doyle/Saunders (1990)	Retail (12 categories) 156w (n.a.)	3	USD	A	D	A	Decreasing LN-shape (A)	Yes	Yes	No (campaign specific)	Yes	OLS No Yes	No	Yes Yes
Dekimpe/Hanssens (1995)	Retail (Home Improvement) 75m (1980-86)	2	const. USD	A	D	A	Flexible / implicitly (T)	Yes	Implicitly	Implicitly	Yes	OLS / VAR Yes Yes	No	No No
Bhargava/Donthu (1999)	Museum 12w (n.a.) L Sports Center 15w (n.a.) 25L	2	USD	A	S	A	No	Yes	No		Yes	OLS / Event Analysis No No	No	No No
Berkowitz/Allaway/D'Souza (2001a)	Retailing (n.a.) 52w (n.a.) 3 Stores L	2	No. of spots/ billboards	A	D	A	No	Yes	Yes	Yes	Yes	Grid Search / OLS Yes No	No	No No
Berkowitz/Allaway/D'Souza (2001b)	Retailing (n.a.) 156w (n.a.) L	2	No. of inserts/ billboards	A	D	A	No	Yes	Yes	Yes	Yes	Grid Search / OLS No No	No	No No
Naik/Raman (2003)	Clothing (Docker Jeans) 47m (1994-97)	2	USD	A	D	A	Decreasing LN-shape (A)	Yes	Yes	No	Yes	Kalman Filter (ML) Yes Yes	No	Yes Yes
Vakratsas/Ma (2005)	Car (2 SUV brands) 123m (1990-00) 108m (1992-00)	3	USD	A	D	A	No	Yes	Implictly	Implicitly	Yes	VAR No No	No	Yes, Scenarios Yes
Dertouzos/Garber (2006)	Military (Armed Services) 36m (1982-84) 66R	4	USD per 1k young men	L	D	A	S-Shape (A)	Yes	Yes	No	Yes	NLS No No	No	No No
Smith/Gopalakrishna/Chatterjee (2006)	Retail (Home Improvement) 104w (2002-03)	4-8 (by stage)	Stage 1: USD Stage 2: Leads Stage 3: Appointment	M	D	A	Decreasing Elasticity (A) Interactions LN-shape (A)	with Elasticity	Yes (Stage 1)	Yes (Stage 1)	Yes	SUR (Stage (1) No Simulation (Stage 1)	No	Yes What-if-Analysis
Naik/Schultz/Srinivasan (2007)	Car (Corolla) 299w (1996-02)	4	USD	A	D	A	Decreasing SQRT-shape (A)	Yes	Yes	No	Yes	OLS / WKF No No	No	No No
Ansari/Mela/Neslin (2008)	Mail Order (Apparell) 48m (1998-02) 500 Cust.	2	Y/N in period	A (M)	D	I	Flexible / implicitly (T)	Yes	Implicitly	Implicitly	Yes	Bayes (MCMC) No Yes	No	No No
Present Study	Softdrink (n.a.) 251w (2000-05) Germany	4	GRP/ USD	A	D	A	Decreasing (flexible) Tested	Yes	Yes	Yes	No	Kalman Filter (ML) Yes Yes	Yes	Yes Yes

3. Modeling Differential Diminishing Returns

To disentangle the three differential effects across advertising media, we extend the Nerlove-Arrow (NA) goodwill formation model by distinctedly incorporating media-specific diminishing returns, ad effectiveness and carryovers.

We start with extending the Nerlove-Arrow (NA) goodwill formation model for each media i:

$$G_{it} = \beta_i\, g_i\left(u_{it}\right) + \lambda_i\, G_{it-1} + \varepsilon_{it}\,, \tag{1}$$

with G_{it} as the unobserved associated medium-specific Goodwill in week t for all media i (i=1,…, N), β_i and λ_i for the medium-specific ad effectiveness and carryover respectively, u_{it} as the gross rating points of medium i in week t, and $g_i(\bullet)$ as the medium-specific diminishing returns function. ε_{it} is a medium-specific error term that follows the normal distribution, $N(0, \sigma^2_{\varepsilon_i})$.

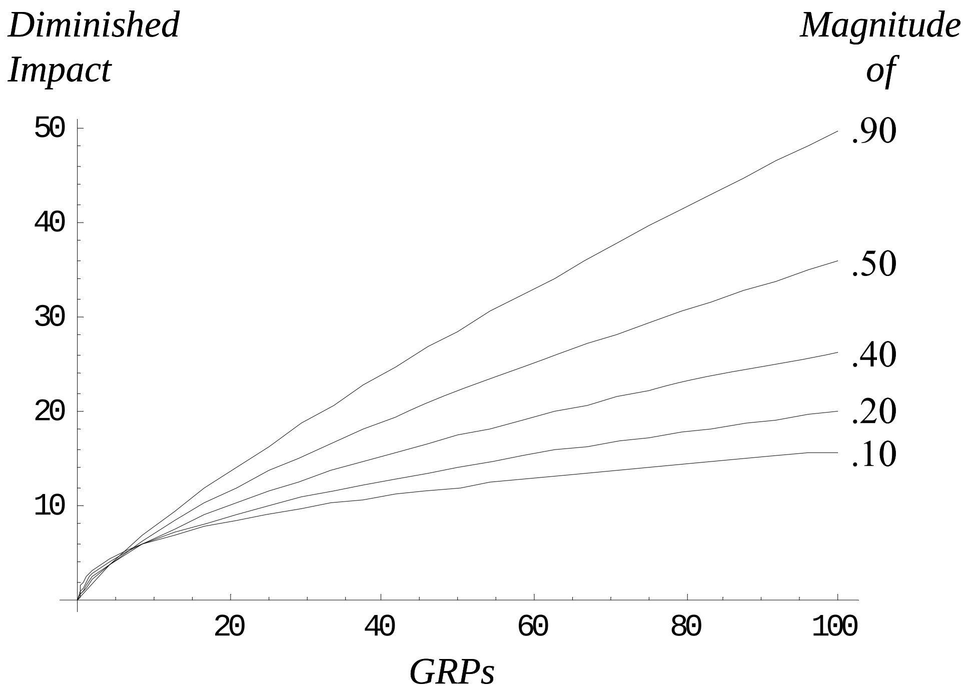
In previous studies, the diminishing returns function is specified to follow either a square-root or logarithmic pattern, i.e., $g_i(u) = \sqrt{u}$ or Ln(u). Consequently, *every* medium i had the *same* diminishing returns, by construction. In contrast, to relax this assumption, we specify the diminishing returns function by the family of shapes,

$$g_i(u) = \frac{u^{\alpha_i}}{\alpha_i}\,, \tag{2}$$

which are indexed by the parameter α_i, $0 < \alpha_i < 1$, representing the *degree* of diminishing returns for the medium i.

Figure 2 illustrates this concept of the degree of diminishing returns. By varying the magnitude of α from large to small, we track the color band from top to bottom to observe the associated shape of the diminishing returns function. As the magnitude of α decreases, the returns of an additional unit diminish faster. When $\alpha = 1$, g(u) = u, we get the extreme case of no diminishing returns. On the other hand, as α tends to 0, $g(u) = \lim_{\alpha \to 0} \left(\frac{u^{\alpha}}{\alpha} \right) = Ln(u)$ reveals the limiting case of strong diminishing returns. In between these two extremes, when $\alpha \in (0, 1)$, Figure 2 shows varying *degrees* of diminishing returns. For example, when $\alpha = 0.5$, we obtain $g(u; \alpha = 0.5) = 2\sqrt{u}$, which marks the intermediate case of diminishing returns used in the literature.

Figure 2. Differential Diminishing Returns



Although this variable family of shapes does not include an s-shaped response, it serves our purpose of testing for the existence of differential diminishing returns. Additionally, as Vakratsas et al. (2004) point out, an s-shape response is unlikely to occur in mature product categories like the one that is studied here.

We note that managers do not directly observe the medium-specific goodwill G_{it} in equation (1) and so they commission a market research firm to measure medium-specific awareness levels A_{it} prevailing in the market place for each week t. Hence, the observed awareness level serves as the fallible proxy, i.e.,

$$A_{it} = G_{it} + \omega_{it}, \tag{3}$$

where error term $\omega_{it} \sim N\left(0, \sigma^2_{\omega_i}\right)$. Later we define equations (1) through (3) as a state-space model linking observed medium-specific awareness levels to GRPs with the respective error terms.

To summarize, the preceding model separates the role of three different sources of media impact: ad effectiveness and carryover of advertising in each medium i (β_i and λ_i in equation (1)) as well as the differential diminishing returns of GRPs in every medium i (α_i in equation (2)). We next propose an approach to estimate these parameters using market data.

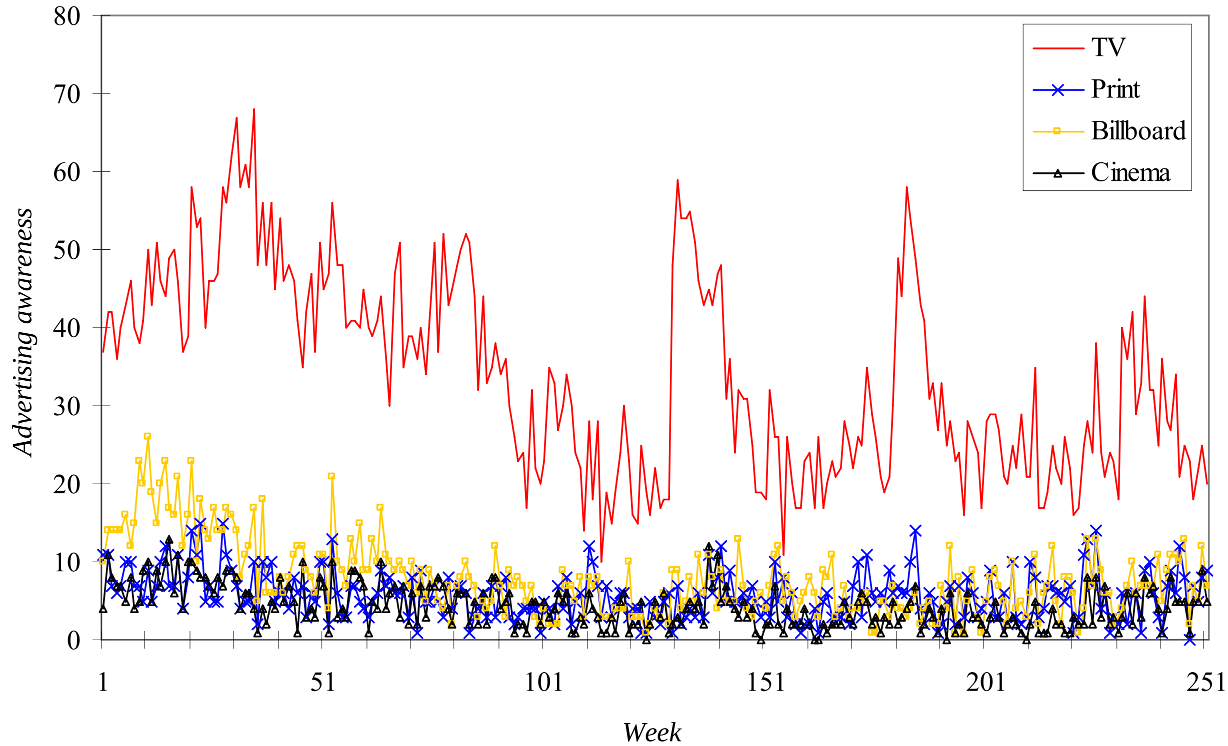
4. Data & Estimation

4.1. Data Description

The empirical application is based on the German soft drink market with two dominating brands, where the bigger brand spends 165 million Euros over five years on four different media. The smaller brand free rides with a total spending of less than 0.5 million Euros. Prices per standard unit for both dominating brands have been stable over the timespan considered. Media specific awareness levels have been collected through weekly surveys with consumers and provided by a leading advertising research agency over 251 weeks from 2000 to 2005. These scores are available to the company approximately 4-6 weeks after the commissioning of the data. Figure 3 shows the dynamics of the awareness levels by week. We use unsupported advertising awareness because this measure is collected by medium as compared to (un-) supported brand awareness and (un-) supported advertising resonance.

Medium-specific advertising awareness has been built up over time through the exposure of consumers to advertising, measured in gross rating points (GRPs). Figure 4 presents the GRPs bought by medium over time. As can be infered from the comparison of figure 3 and 4, GRPs in different media seemingly translate into awareness levels differently, i.e., billboard GRPs are quite high compared to TV, but awareness for TV is much higher than for billboards.

Figure 3. Unsupported Awareness by Medium



These GRPs are bought by the media agency 3-12 months in advance with TV stations, publishing companies of newspapers and magazines (here summed up as print media), billboard agencies and cinema chains. As awareness levels measured are available to the media agency only 4-6 weeks after airing, the decision to spend on GRPs in a given week and medium is not based on observed awareness levels. Hence, endogeneity is not encountered here.

Figure 4. GRPs by Medium

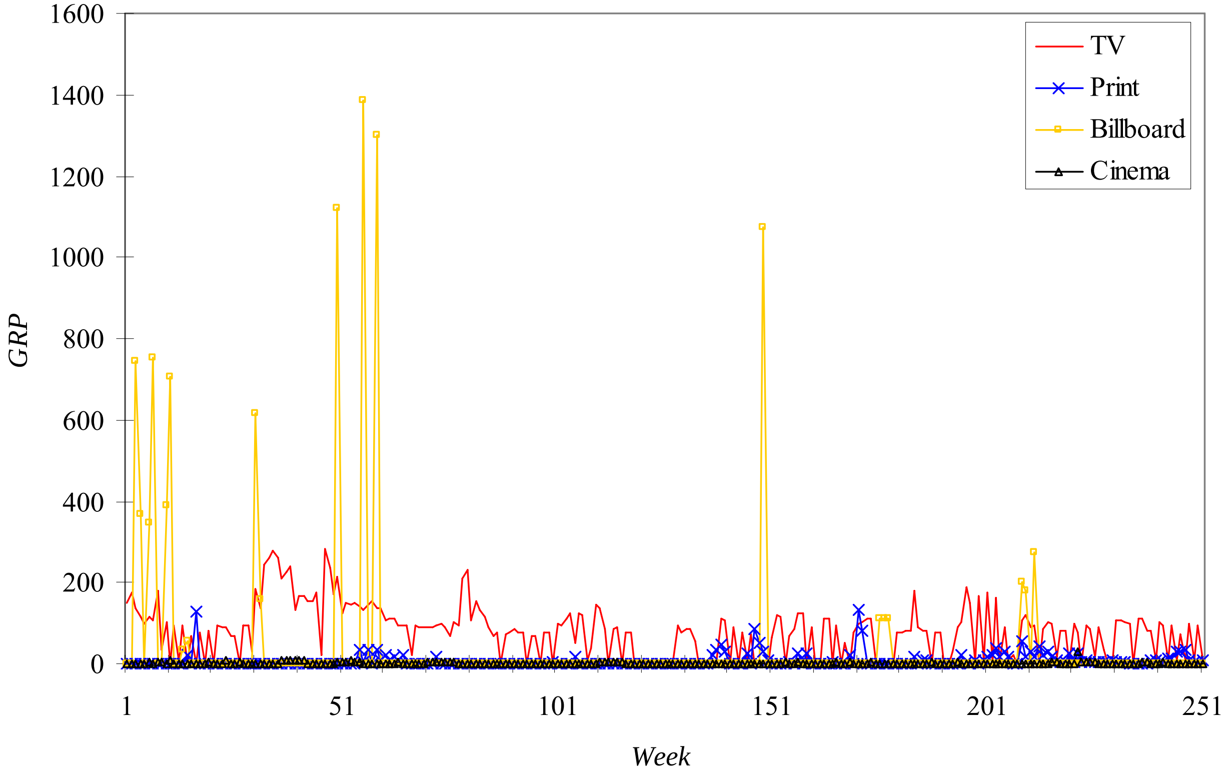


Table 2 summarizes the descriptives across medium-specific awareness levels and GRPs. The data supports the notion that the dynamic relationships across measures and media are non-trivial and obviously differing. To capture the varying media dynamics and simultaneously disentangle these medium-specific effects, we need to develop an appropriate estimation approach which we will outline in the next section.

Table 2. Descriptive Statistics

Statistics		N	Mean	Std. Dev.	Minimum	Maximum
Advertising Awareness+	TV	251	34.02	12.16	10	68
	Print (PR)	251	5.78	3.03	0	15
	Billboard (BB)	251	7.86	4.77	1	26
	Cinema (CI)	251	4.33	2.73	0	13
GRP*	TV	251	79.65	63.59	0	283
	Print (PR)	251	7.15	16.94	0	133
	Billboard (BB)	251	40.00	181.53	0	1,387
	Cinema (CI)	251	1.19	2.73	0	30

+ Awareness by media   * GRP - Gross Rating Points

4.2. Estimation Approach

To estimate parameters of (1), (2) and (3), we use the standard state-space form (see, e.g., Harvey 1994, Ch. 3). We express equations (1) and (2), together, as the transition equation (5) and equation (3) as the observation equation (4), linking GRPs to the observed medium-specific awareness levels A_{it}:

$$\begin{bmatrix} A_{1t} \\ A_{2t} \\ A_{3t} \\ A_{4t} \end{bmatrix} = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{bmatrix} \begin{bmatrix} G_{1t} \\ G_{2t} \\ G_{3t} \\ G_{4t} \end{bmatrix} + \begin{bmatrix} \omega_{1t} \\ \omega_{2t} \\ \omega_{3t} \\ \omega_{4t} \end{bmatrix} \tag{4}$$

$$\begin{bmatrix} G_{1t} \\ G_{2t} \\ G_{3t} \\ G_{4t} \end{bmatrix} = \begin{bmatrix} \tilde{\lambda}_1 & & & 0 \\ & \tilde{\lambda}_2 & & \\ & & \tilde{\lambda}_3 & \\ 0 & & & \tilde{\lambda}_4 \end{bmatrix} \begin{bmatrix} G_{1t-1} \\ G_{2t-1} \\ G_{3t-1} \\ G_{4t-1} \end{bmatrix} + \begin{bmatrix} \beta_1 & & & 0 \\ & \beta_2 & & \\ & & \beta_3 & \\ 0 & & & \beta_4 \end{bmatrix} \begin{bmatrix} (u_{1t})^{\tilde{\alpha}_1} / \tilde{\alpha}_1 \\ (u_{2t})^{\tilde{\alpha}_2} / \tilde{\alpha}_2 \\ (u_{3t})^{\tilde{\alpha}_3} / \tilde{\alpha}_3 \\ (u_{4t})^{\tilde{\alpha}_4} / \tilde{\alpha}_4 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{bmatrix} \tag{5}$$

Then we apply the Kalman filter (e.g., Naik et al. 1998; Bass et al. 2005) to determine recursively (i.e., for each t = 1,...,T) the mean and covariance of the unobserved medium-specific goodwills G_{it}, given the observed awareness sequence up to the previous period, $H_{t-1} = \{\mathbf{A}_1, ..., \mathbf{A}_{t-1}\}$, where $\mathbf{A}_t$ is a vector of A_{it}. Based on those means and covariances of G_{it}, we compute the log-likelihood of observing the awareness sequence $\{A_1, A_2, \ldots, A_T\}$, which is given by

$$\begin{aligned} L(\Theta; H_T) = \; & Ln(p(A_1, A_2, \cdots, A_T; \Theta)) \\ & = Ln(p(A_T \mid H_{T-1}; \Theta)) + Ln(p(A_1, A_2, \cdots, A_{T-1}; \Theta)) \\ & = Ln(p(A_T \mid H_{T-1}; \Theta)) + Ln(p(A_{T-1} \mid H_{T-2}; \Theta)) + Ln(p(A_1, A_2, \cdots, A_{T-2}; \Theta)) \\ & \vdots \\ & = \sum_{t=1}^{T} Ln(p(A_t \mid H_{t-1}; \Theta)), \end{aligned} \tag{6}$$

where p(· | ·) is the conditional density of awareness $\mathbf{A}_t$ given the information history H_{t-1}, and Θ is a K x 1 parameter vector.

The composition of vector Θ in an empirical application with four types of media is as follows. Besides the four initial medium-specific goodwills G_{i0} and respective noise terms ($\sigma_{i\varepsilon}$, $\sigma_{i\omega}$), the parameters of interest are the ad effectiveness coefficients (β_1, β_2, β_3, β_4), the carryover effects (λ_1, λ_2, λ_3, λ_4), and the degrees of diminishing returns (α_1, α_2, α_3, α_4). To constrain the latter subset γ = (λ_1, λ_2, λ_3, λ_4, α_1, α_2, α_3, α_4) within the unit interval $\gamma_j \in (0,1)$, we apply the transformation

$$\gamma_j = \exp(\tilde{\gamma}_j)/(1 + \exp(\tilde{\gamma}_j)) \tag{7}$$

where $\tilde{\gamma} = (\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3, \tilde{\lambda}_4, \tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4)$ is an unconstrained vector taking values in $\Re^8$ and *j* denotes the corresponding element of the vectors γ and $\tilde{\gamma}$. This transformation ensures that the magnitude of every element of γ_j lies within the unit interval (0, 1) for any value of $\tilde{\gamma}_j \in (-\infty, \infty)$. Thus $\Theta = (G_{i0}, \beta_i, \sigma_{i\varepsilon}, \sigma_{i\omega}, \tilde{\lambda}_i, \tilde{\alpha}_i)'$ (i=1,2,3,4) is the parameter vector to be estimated (compare notation in our state-space form in equation (5)).

To obtain the parameter estimates, we next maximize (6) with respect to Θ:

$$\hat{\theta} = ArgMax\; L(\Theta) \tag{8}$$

Finally, to assess significance of the estimated parameters, we can obtain the standard errors by taking the square-root of the diagonal elements of the inverse of the matrix:

$$\hat{J} = \left[-\frac{\partial^2 L(\Theta)}{\partial \Theta \partial \Theta'} \right]_{\Theta=\hat{\theta}} \tag{9}$$

where the Hessian of $L(\Theta)$ is evaluated at the estimated values $\hat{\theta}$.

Robust Inferences

However, our models can be potentially mis-specified. Hence, we seek to make statistical inferences robust to misspecification errors. To this end, we conduct Huber-White robust inferences (see White 1982) by computing the *sandwich estimator*,

$$Var(\hat{\theta}) = \hat{J}^{-1} \hat{V} \hat{J}^{-1} \tag{10}$$

where V is a K × K matrix of the gradients of the log-likelihood function; that is, $V = P'P$, and P is a T × K matrix obtained by stacking the 1 × K vector of the gradient of the log-likelihood function in (6) for each of the T observations. In correctly specified models, $J = V$ and so both the equations (9) and (10) yield exactly the same standard errors (as they should); otherwise, for mis-specified models, we use the robust standard errors given by the square-root of the diagonal elements of the matrix in equation (10).

5. Empirical Results

Estimating the model with the developed estimation procedure for N=251 periods yields the results shown in Table 3. Before discussing the empirical results we comment on the diagnostics briefly.

5.1. Fit, Log-Likelihood, and Diagnostics

Fit and Log-likelihood

Maximizing the Log-likelihood yields a value of -1,747. The corresponding values for AIC, AICc and BIC are given in Table 3 and suggest a satisfactory model fit. The observation noise is smaller than awareness standard deviations (compare with Table 2), i.e., a substantial fraction of awareness variation can be explained by the model. The transition noise shows that the modeled effects capture goodwill dynamics to a satisfactory degree. Additionally, the initial goodwill values are all in plausible ranges and vary substantially by media.

Diagnostics

Following Harvey (1994, p. 256), we conduct diagnostics based on the residuals defined by $\tilde{v}_t = (Y_t - za_{t|t-1}) / \sqrt{f_t}$, where f_t = Var(Y_t) for t = 1, …, T. Specifically, we test for serial correlation, heteroscadesticity and parameter constancy.

Serial Correlation. The sample auto-correlations for awareness levels in TV, Print, Billboard and Cinema are (-0.04, 0.06, -0.02, -0.05). To test whether they significantly differ from zero, we use the Box-Ljung statistic

$$Q = \frac{T(T+2)}{T-1}\left[\frac{\sum_{t=2}^{T}(\tilde{v}_t - \bar{\bar{v}})(\tilde{v}_{t-1} - \bar{\bar{v}})}{\sum_{t=1}^{T}(\tilde{v}_t - \bar{\bar{v}})^2}\right]^2, \tag{11}$$

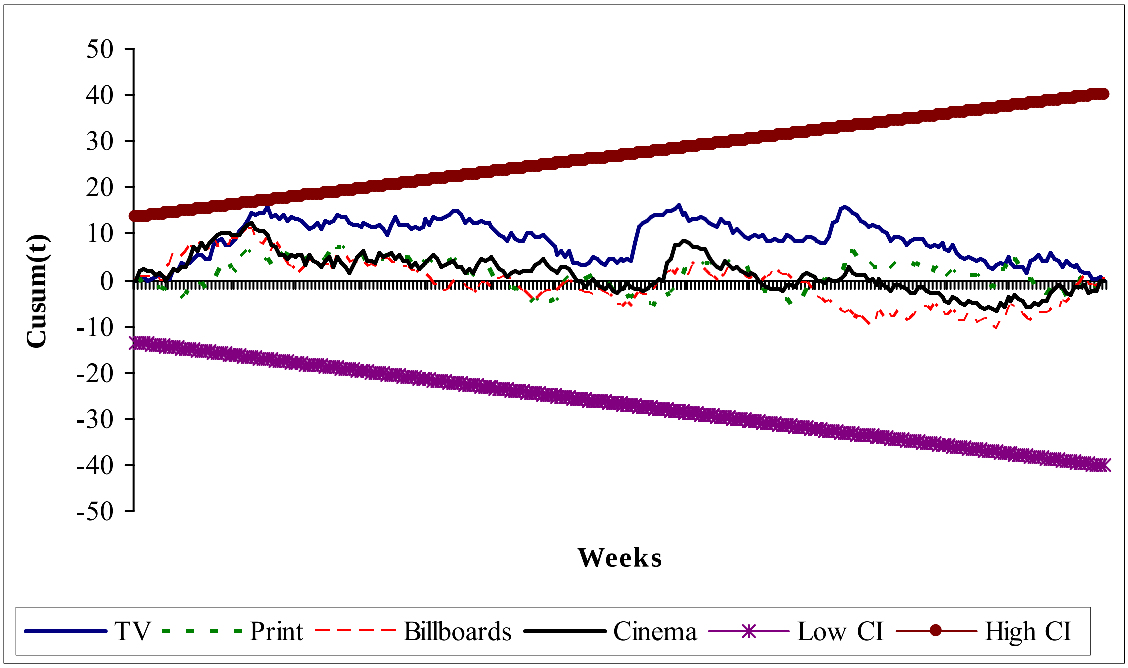
which is distributed as χ^2 with 1 degree of freedom (Harvey 1994, p. 259). For the sample, Q = (0.37, 0.98, 0.06, 0.60), which do not exceed the critical χ^2 value of 3.84 at the 95% confidence level. Hence, serial correlations are statistically negligible.

Heteroscadesticity. A test statistic for heteroscadesticity is

$$H(h) = \frac{\sum_{t=T-h+1}^{T} \tilde{v}_t^2}{\sum_{t=1}^{1+h} \tilde{v}_t^2}, \tag{12}$$

where h = T/3, and H(h) is distributed as F(h, h) with degrees of freedom h (see Harvey 1994, p. 259). For this sample, h = 251/3 ≈ 84 and H = (0.96, 1.29, 0.77, 0.63), which do not exceed the critical value F(84,84) = 1.435 at the 95% confidence level. Hence, the residuals do not exhibit heteroscadesticity.

Figure 5. Cumulative Residuals



Parameter Constancy. We use the test due to Brown, Durbin and Evans (1975), who note that the residuals would depart over time if the true parameters evolved while the estimated parameters were considered constant. Hence, to test parameter constancy, we compute the cumulative residuals, CUSUM, defined as

$$\text{CUSUM}(t) = \hat{\sigma}^{-1} \sum_{j=1}^{t} \tilde{v}_j, \quad j = 1, \ldots, T. \tag{13}$$

Then, if the above statistic lies within the significance lines given by $\pm [a\sqrt{T} + 2at/\sqrt{T}]$, where a = 0.845 for 10% level of significance, the model parameters are considered constant. For each medium, we compute the CUSUM statistic and present the results in the Figure 5 above. We observe that the cumulative residuals stay within the confidence bands. Hence, model parameters are stable over time.

Table 3. Differential Advertising Effects

Parameter	*Medium*	*Value*	*Estimate*	*Std. Error (Robust Inf.)*	*t-value (Robust Inf.)*
Panel A: Ad Effectiveness					
	TV, β_1		0.3197	0.1146	2.79
	Print, β_2		0.0205	0.0053	3.86
	Outdoor, β_3		0.0628	0.0183	3.44
	Cinema, β_4		0.0639	0.0302	2.12
Panel B: Carryover					
Estimates	TV, $\tilde{\lambda}_1$		2.1909	0.3662	5.98
	Print, $\tilde{\lambda}_2$		3.8280	0.3968	9.65
	Outdoor, $\tilde{\lambda}_3$		3.2018	0.2790	11.48
	Cinema, $\tilde{\lambda}_4$		3.2853	0.5458	6.02
Retransformed Values	TV, λ_1		0.8994	0.0331	27.15
	Print, λ_2		0.9787	0.0083	118.37
	Outdoor, λ_3		0.9609	0.0105	91.67
	Cinema, λ_4		0.9639	0.0190	50.78
Panel C: Diminishing Returns					
Estimates	TV, $\tilde{\alpha}_1$		-1.1557	0.5121	-2.26
	Print, $\tilde{\alpha}_2$		-1.0766	0.8851	-1.22
	Outdoor, $\tilde{\alpha}_3$		-0.6328	0.2728	-2.32
	Cinema, $\tilde{\alpha}_4$		1.0054	2.5468	0.39
Retransformed Values	TV, α_1	0.2394	0.2394	0.0933	2.57
	Print, α_2*	0.5000	0.2541	0.1678	1.51
	Outdoor, α_3	0.3469	0.3469	0.0618	5.61
	Cinema, α_4*	0.5000	0.7321	0.4995	1.47
Panel D: Initial Goodwill, Observation and Transition Noise					
Initial Goodwill	TV, G_{10}		37.6291	2.1069	17.86
	Print, G_{20}		9.7398	2.4736	3.94
	Outdoor, G_{30}		13.9580	4.1913	3.33
	Cinema, G_{40}		8.0658	5.3130	1.52
Transition Noise	TV, $\sigma_{\varepsilon 1}$		4.2627	0.6030	7.07
	Print, $\sigma_{\varepsilon 2}$		0.0000	0.0130	0.00
	Outdoor, $\sigma_{\varepsilon 3}$		0.4254	0.1283	3.32
	Cinema, $\sigma_{\varepsilon 4}$		0.3355	0.1555	2.16
Observation Noise	TV, $\sigma_{\omega 1}$		4.2455	0.4043	10.50
	Print, $\sigma_{\omega 2}$		2.8241	0.1341	21.05
	Outdoor, $\sigma_{\omega 3}$		3.0265	0.1707	17.73
	Cinema, $\sigma_{\omega 4}$		2.1653	0.1133	19.12
Maximized Log-Likelihood, L*				-1,747	
AIC / AICc / BIC			3541.43	4106.98	3626.04

* As estimate is not significant it is set to zero and thus transformes into 0.5 or SQRT-fct.

5.2. Ad Effectiveness (Panel A)

All media exhibit ad effectiveness to a varying degree. Here, TV is by far the most effective medium. In this context, the emotional spots put on air by the soft drink brand may have been working best in TV. Print is the least effective medium, but inspecting the copies that were put into magazines during this period resembled mostly plain images drawn from the TV spot. As print media are known to have advantages with information oriented messages, the emotional print copies used seem to have been insufficiently adapted to the medium. With respect to below the line media, i.e., cinema and outdoor, the copies have been specifically designed for the respective medium. Cinema carries similar advantages like TV as a medium, so the emotional spots created work accordingly. For outdoor, again distinct copies have been created that nevertheless connected to the TV campaign. Overall, these two below the line media have been equally effective.

5.3. Differential Carry-Over (Panel B)

All Carryover parameters are significant at high levels indicating that a dynamic model is necessary. Please note that Panel B in Table 3 contains two parameter sets. The first represents the estimates from the Kalman Filter procedure, whereas the second set contains the retransformed estimates from $\tilde{\gamma}_j \in (-\infty, \infty)$ to γ_j as well as the correspondingly converted standard errors and t-values from the robust inference procedure. In this context, carryover for TV is lowest and most different from other media. Hence, assuming a single carryover across media may distort effectiveness and differential diminishing return parameters for other media. Although print media have a comparatively low effectiveness, they exhibit the highest carryover. This supports the notion that print media gain the highest attention by the consumer when compared to TV, cinema and outdoor. So once readers pay attention to the copy, they remember them better compared to the other media (cf. section 2). Cinema and outdoor media show only slightly lower carryovers. For cinema, the captive environment and event character may enhance long term effectiveness with consumers (e.g., Ewing, Plessis and Foster 2001). Concerning outdoor advertising, the GRPs have been spent quite focused within comparatively short time frames achieving a perceived ubiquity temporarily. According to the agency, many people remembered the campaigns quite long for this perceived ubiquity. Deriving the associated long term impact of these carryover values we find that they correspond from 3 months (ca. 10 weeks) for TV to approximately 10 months (ca. 46 weeks) for print. These values add further evidence to the value ranges compiled by, e.g., Leone (1995) and Lodish et al. (1995), who estimate regular long term impacts to range from 6-9 months on average.

5.4. Differential Diminishing Returns (Panel C)

We find support for the presence of differential diminishing returns. Like for carryovers, Table 3 contains two corresponding parameter sets, i.e., one from the Kalman Filter estimation and the similarly converted parameters. Panel C shows that TV and outdoor returns diminish slower than LN(*), but faster than SQRT(*). For the other two media, the parameter estimates are not significantly different from zero, which - when converted for $\tilde{\gamma}_j = 0$ to γ_j - would correspond to diminishing returns that follow the SQRT(*)-function.

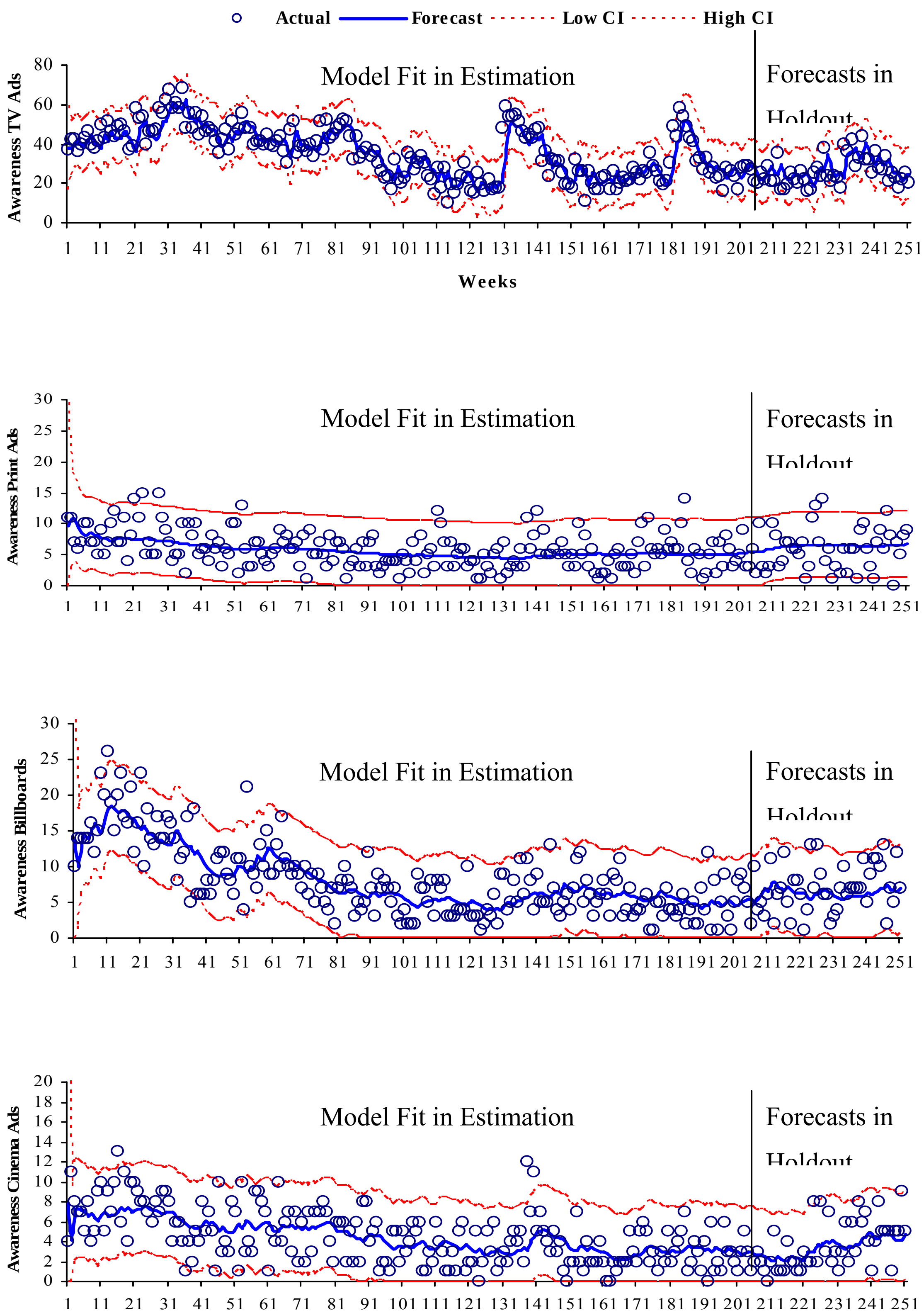
But are these effects jointly and significantly different from zero, i.e., due to the value transformation different from SQRT(*)? To address this issue we test the joint null hypotheses H_0: $\tilde{\alpha}_1 = \tilde{\alpha}_2 = \tilde{\alpha}_3 = \tilde{\alpha}_4 = 0$. To this end, we apply the form $R\tilde{\alpha} = r$, where $R = \mathbf{I}_{4x4}$, $\tilde{\alpha} = \left(\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4\right)'$, and $r = \mathbf{0}_{4x1}$. We evaluate $\Sigma = \mathrm{Cov}(R\tilde{\alpha}) = R\Sigma_{\tilde{\alpha}}R'$ and compute the statistic $T = (R\tilde{\alpha} - r)' \Sigma^{-1} (R\tilde{\alpha} - r)$, which is distributed χ^2 with degrees of freedom equal to the rank(R). For our data, T = 11.27 and the critical χ^2 value = 9.49 for 4 degrees of freedom at 95% level. Because T exceeds the critical value, we reject the null hypothesis, indicating that (i) $\tilde{\alpha}_i$ are not equal across media (which lends support to the managerial belief that differential diminishing returns prevail), and that (ii) neither SQRT(*) nor LN(*) is the appropriate diminishing returns function for at least two media within this application.

5.5. Cross Validation

We report the *out-of-sample* forecasting performance of the model. To this end, we re-estimate the model using the estimation sample consisting of data from the first four years (208 weeks) and then predict the awareness levels for the subsequent 43 weeks (over 9 months).

In Figure 6 below, we present both the model fit in the estimation sample and one-week-ahead forecasts in the holdout period for TV, print, billboards and cinema. In this figure, we observe three points; first, in-sample model fit (in the first 208 weeks) and out-of-sample forecasts (in the last 43 weeks) indicate comparable accuracy. Second, the model tracks TV

Figure 6. One-step-ahead Predictive Forecasts in Holdout Period



awareness levels remarkably well, including the turning points, because TV responds to managerial actions better than the other three media (see Table 1). Finally, for all four media, the observed awareness levels are well within the 95% forecast intervals. Overall, the out-of-sample model performance seems satisfactory.

6. Marketing Implications

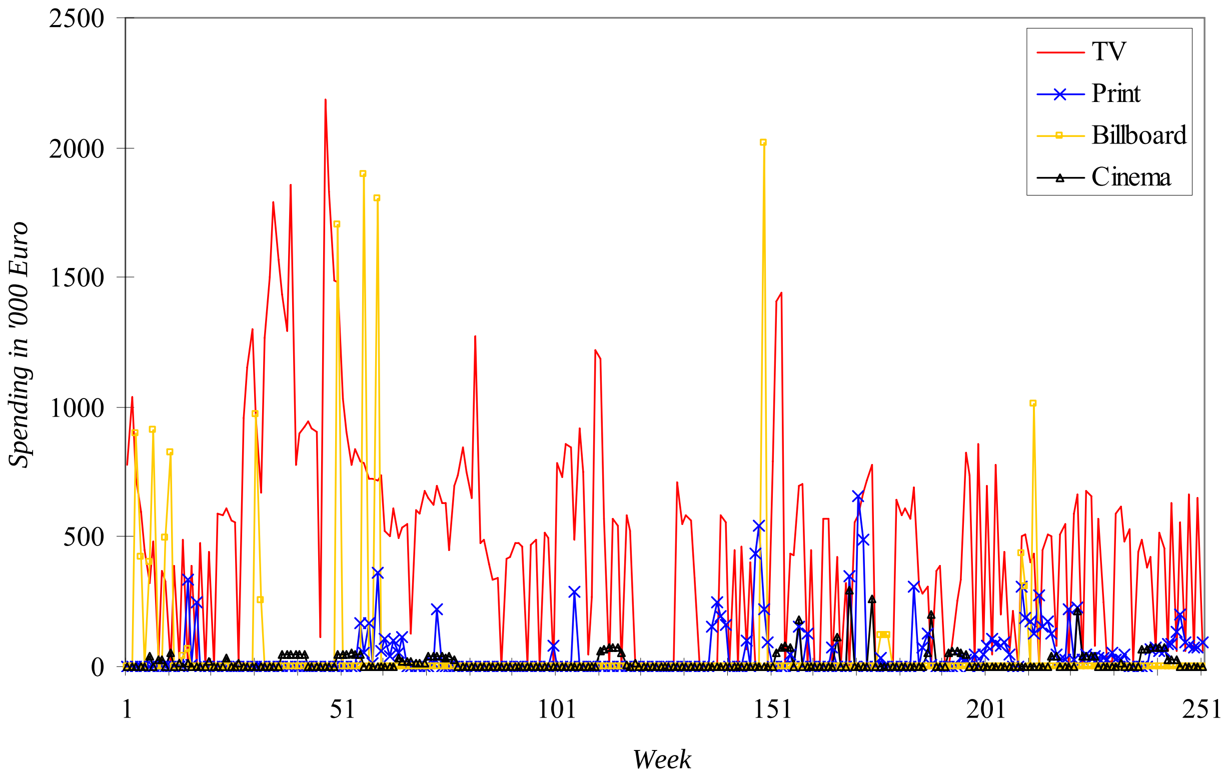
The empirical application confirms the existence of differential diminishing returns, ad effectiveness, and carryovers. But what implications does their existence have on the optimal allocation of an advertising budget? To assess this impact, we have to connect GRPs with the media budgets spent. As with the above mentioned differential effects, the same amount of dollars may buy different quantities of GRPs across different media. Accordingly, we start with accounting and testing for differential media buying efficiencies. After establishing the presence of differential media buying efficiencies, we followingly integrate this effect into the model through the specification of a joint profit function. This profit function will then provide the basis for deriving the optimal allocations rules under four differential effects across media.

6.1. Differential Media Buying Efficiencies

Media buying efficiency has been of continuing interest since the the early days of advertising. Duffy (1938) is an early source discussing huge variations in the cost-per-thousand measure across several advertising media of the time, such as print, radio and outdoor. His investigation inspired many studies concerning the efficiency of advertising across different media. A representative and recent study is that of Luo and Donthu (2001). They investigate the efficiency of 63 major advertising companies in the U.S. Their main conclusion is that the majority of advertisers allocates their budgets suboptimal, thus confirming that media buying efficiency is still an important issue. As their findings are based on a cross-sectional DEA analysis, they call followingly for a longitudinal perspective of advertising efficiency in multimedia settings.

Across all studies reviewed here (cf. Table 1), only Hugues (1975) investigates media efficiency. All other studies do not expilictly integrate media buying efficiency into their modeling framework, although most of them measure the advertising input in USD. Accordingly, media efficiency may be implicitly accounted for, but is at the same time at least entangled with ad effectiveness. Hence, there is gap in the literature with respect to accounting for differential media buying efficiencies while simultaneously separating the three other sources of media effectiveness.

Figure 7. Media budgets by medium



According to Danaher and Rust (1994), modeling advertising efficiency requires a mathematical function which relates costs to GRPs obtained that - due to discounting with more GRPs bought - exhibits decreasing unit costs with additional GRPs acquired. We follow their recommendation and establish the notion of media buying efficiencies in a simple linear function

$$v_{it} = a_i + b_i \, u_{it}, \tag{14}$$

for each medium $i = 1, ..., 4$. In equation (14), v_{it} represents the monetary cost of buying u_{it} GRPs in the medium i in week t. Furthermore, the parameters a_i and b_i, respectively, estimate the *minimum cost* of buying GRPs and the *marginal cost* per incremental GRP in a medium i.

Table 4. Description on Media Spendings

Statistics		N	Mean	Std. Dev.	Minimum	Maximum
Spending*	TV	251	480.33	408.58	0	2,188
	Print (PR)	251	45.42	95.95	0	654
	Billboard (BB)	251	59.12	269.96	0	2,020
	Cinema (CI)	251	17.12	38.73	0	296

* Spending in 1,000 Euros

Figure 7 illustrates the spending data across the four media employed in our soft drink case. On average across the 251 weeks, the company spent 480.33 thousand Euros on TV (Std. Dev. 408.58), 45.42 thousand Euros (95.95) on print, 59.12 thousand Euros (269.96) on billboards, and 17.12 thousand Euros (38.72) on cinema advertising. Comparing figures 7 and 4, it is obvious that the GRPs incur different costs across media. Hence, media buying efficiency seems to vary with media, which seems to confirm the issues raised already by Duffy (1938).

Estimating (14) for all four media employed here yields the results presented in Table 4. Accordingly, media buying efficiency differs across media with TV exhibiting the highest minimum cost, but medium marginal costs. Print requires relatively high minimum costs for an initial GRP point, with comparable marginal costs. Both below the line media have lower minimum costs, but different marginal cost. Cinema actually is the most expensive medium in terms of marginal costs. The varying efficiencies across media lead to medium-specific cost curves, hence media co-exist.

Table 5. Media Efficiencies

Medium	*Minimum Cost (€)*	*Marginal Cost (€)*	*adj. R^2*
TV	23,674 *(1.27)*	5,734 *(31.32)*	0.796
Print	10,752 *(3.17)*	4,850 *(26.25)*	0.733
Outdoor	925 *(0.26)*	1,455 *(74.77)*	0.957
Cinema	5,737 *(2.91)*	9,558 *(14.40)*	0.452

t-values in parantheses

At this point, we have shown that there is gap in the literature not only concerning disentangling the three previously addressed issues of differential media effectiveness, but that also differential media buying efficiencies have not been jointly investigated with the identified multimedia studies. Following Danaher and Rust (1994), we show that media buying efficiencies - like ad effectiveness, carryovers, and diminishing returns - indeed vary across different media, too. In the next chapter, we will extend our modeling framework to account for varying media buying efficiencies in order to infer normative implications for optimizing multimedia budgets.

6.2. Cross Media Allocation - Normative Implications

Using the awareness data by media, managers want to determine the amount of money to be spent (i.e., budget) and allocate that budget across multiple media so that they maximize the total expected awareness. Building on our framework from section 3.1, we express the objective function as total expected awareness net of total money spent as follows (e.g., Naik et al. 1998, Bass et al. 2007):

$$\underset{u_i}{\text{Max}}\ J(u_1, u_2, .., u_n),\ \text{where}\ J(u_1, u_2, .., u_n) = \int_{t=0}^{\infty} e^{-\rho t} (\sum_{i=1}^{i=n} \left(mE[A_{it}] - \nu_{it}\right) dt \qquad (15)$$

where, $A_{it} = G_{it} + \omega_{it}$, $\omega_{it} \sim N[0, \sigma^2_{\omega_i}]$, $\nu_{it} = a_i + b_i u_{it}$, m and ρ representing the profit impact per awareness point and discount rate respectively, subject to the state equation

$$dG_{it} = -\beta_i g_i(u_{it}) + \lambda_i G_{it-1} + \varepsilon_{it} \qquad (16)$$

which, in continuous time, becomes

$$\frac{dG_i(t)}{dt} = -\beta_i g_i(u_{it}) + \lambda_i G_i(t) \qquad (17)$$

where we denote the functional dependence of G_i on t by the usual notation $G_i(t)$ and we ignore the error term ε_{it} in our deterministic formulation of the optimal control problem.

Formally, the objective is to find the best possible investment plans across all N media, $\{u_1^*, u_2^*, \cdots, u_N^*\}$, that yields the largest value of $J(u_1, u_2, \cdots, u_N)$ subject to the model equations (1), (2) and (3), while accounting for varying media buying efficiencies according to (14). In other words, this framework separates the role of three different sources of media effectiveness: ad effectiveness and carryover of advertising in each medium i (β_i and λ_i in equation (1)), the differential diminishing returns of GRPs in every medium i (α_i in equation (2)), while simultaneously accounting for the buying efficiency across various media i (a_i and b_i in equation (14)). Based on this holistic framework, we investigate the normative implications for optimal budget allocations; for example, if diminishing returns parameter α increases, should managers increase or decrease the budget allocated to that medium?

To gain such insights, we apply optimal control theory to solve the allocation problem formulated in the previous section and then conduct comparative statics analyses (e.g., Kamien and Schwartz 1991, Sethi and Thompson 2000). We present the optimal allocation strategy in the following proposition.[1]

Proposition 1. Maximizing (15) subject to (1), (2), (3) and (14), the optimal allocation across N media is given by

$$u_{i,(2)}^* = \left(\frac{m\,\beta_i}{b_i(1+\rho-\lambda_i)} \right)^{\left(1/(1-\alpha_i)\right)}, \;\forall\, i=1, 2,\ldots, N. \tag{18}$$

Using Proposition 1, we analyze the comparative statics and present the findings in the next proposition.

[1] Proofs are available from the authors upon request.

Proposition 2. Table 7 summarizes the main effects with respect to the optimal allocation u_i^*.

Table 6. Normative Implications of Main Effects

Increase in Main Effect	Change in u_i*
Profit Impact of Awareness Point, m	↑
Discount Rate, ρ	↓
Ad Effectiveness, β_i	↑
Carryover Effect, λ_i	↑
Marginal Cost of Medium, b_i	↓
Minimum Cost of Medium, a_i	No change

The results with respect to profit impact per awareness point, discount rate, ad effectiveness and carryover corroborate with the extant literature (see e.g., Naik and Raman 2003), even though it is shown here that the condition concerning the carryover even holds on the medium-specific level. The other implications are new and consistent with managerial intuition – the more expensive the medium per marginal GRP (b_i), the smaller the spending. More importantly, we observe that $\partial u_{i(2)}^* / \partial a_i = 0$ for any medium i. This finding reveals that the optimal allocations do *not* depend on the fixed costs of buying GRPs despite managers' concerns such as "television costs more that print or billboards" or as a respondent in a media study of Barwise and Farley (2005) mentioned, "…project on the Internet… [is] cheaper". Finally, we emphasize that these results hold in general for *every* feasible value of the parameters $(m, \rho, \beta_i, \lambda_i, b_i)'$ - not just the estimated parameter values from this data sample of softdrink brand's advertising.

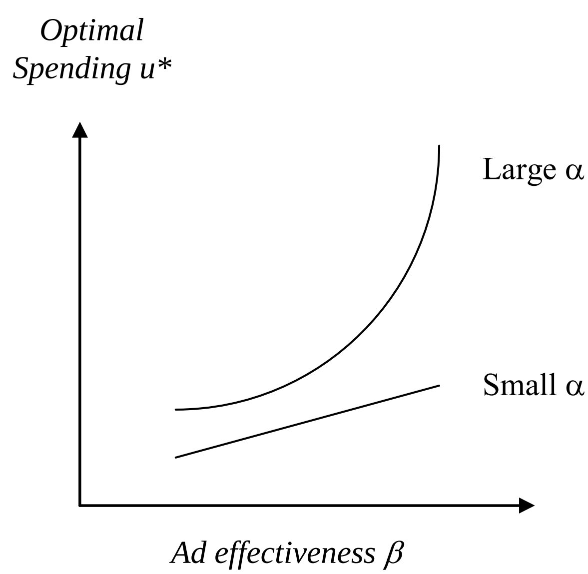
Proposition 3. As α_i increases (i.e., severity of diminishing returns decreases), the optimal allocation $u_{i(2)}^*$ increases, provided the following condition holds:

$$\frac{m\beta_i}{(1+\rho-\lambda_i)} > b_i \tag{19}$$

This condition is not just a technical requirement, but carries substantive meaning: in a given medium i, the discounted long-term ad effectiveness and the profit per awareness point must exceed its marginal cost. Also, this adds to the understanding that in cumulative sales models, the optimal advertising rate is increasing if advertising becomes more efficient (e.g., during the product life cycle; Feichtinger, Hartl, and Sethi 1994; Teng and Thompson 1985).

To better understand the interaction effects between ad effectiveness and diminishing returns for any given medium i, consider the different scenarios of for a small and large value of α illustrated in Figure 8.

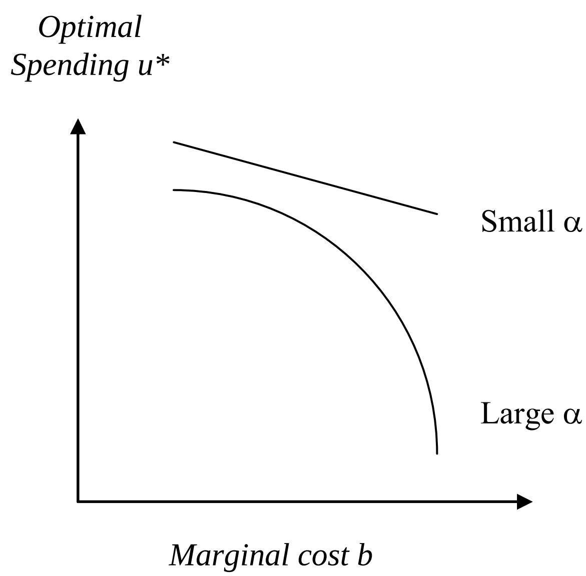
Figure 8. Interaction Effects between Ad Effectiveness and Diminishing Returns



When diminishing returns for a medium are strong, i.e., α is small, then optimal spending u* increases *proportionally* with an increasing ad effectiveness β. In contrast, when dimishing returns are small, i.e., α is large, then optimal spending u* increases *exponentially* with an increasing ad effectiveness β.

Another interesting relationship is found between the marginal cost b and the diminishing returns α of any given medium I, which is shown in Figure 9.

Figure 9. Interaction Effects between Marginal Cost and Diminishing Returns



When diminishing returns are strong, i.e., α is small, then optimal spending for the medium u* decreases *proportionally* with increasing marginal cost b. In contrast, when dimishing returns are small, i.e., α is large, then optimal spending for the medium u* decreases asymptotically with increasing marginal cost b.

More importantly, this proposition sheds light that managers do *not* have to increase spending on a medium as diminishing returns become less severe - note that intuition would suggest otherwise - unless the above condition also holds. Specifically, in the empirical example, the print, outdoor and cinema media are more attractive than TV because of smaller diminishing returns (i.e., $\alpha = 0.5$ each for print and cinema, and outdoor with an $\alpha = 0.35$ versus 0.24 for broadcast; see Table 3). Yet managers spend more on the broadcast media (see Table 2). Why? Because, we suspect, the long-term ad effectiveness of broadcast media is likely to exceed its marginal cost.

7. Conclusions

Our study contributes to the existing literature in several ways. First, we develop an appropriate model that accounts for differential effects across various media, disentangling three confounding sources of media spending impact. Second, we develop a systematic method to estimate the model's parameters and illustrate this approach using market data. We find that the various effects --- ad effectiveness, carryover and diminishing returns --- indeed vary across different media. The results also indicate a first assessment on the general effectiveness of billboard and cinema spendings. Third, we investigate the normative implications for managerial decision making, while simultaneously accounting for differential media buying efficiency. For optimal budgets of a given medium, we show that they also depend on interaction effects of diminishing returns with two other medium specific effects, ad effectiveness and marginal cost. Based on our holistic framework, we demonstrate that managers do *not* have to increase spending on a medium as diminishing returns become less severe, unless discounted long-term ad effectiveness and profit per awareness point exceed its marginal cost, while intuition would suggest otherwise. Hence, with the increasing importance of integrated media campaigns, managers should follow their presumption and account for differential effects across media. The proposed framework should provide them with a toolkit for an improved media-specific budget allocation.

With respect to future research we suggest three avenues. First, managers could use recursive updates of the parameter estimates using incremental data for 4-weeks to adjust their short-term spending allocation in case of substantial changes. Even though we found our parameter estimates to be stable over time, this need not be the general case. Second, our findings of differential media effectiveness should be validated across other brands and categories. Third, competitive advertising might influence differential effects over time resulting in varying parameter estimates that depend not only on one's own advertising implementation.

References

Aaker, D.A. 1975. ADMOD: An advertising decision model. *J. Marketing Res.* **12**(1) 37-45.

Ansari, A., C.F. Mela, S.A. Neslin. 2008. Customer Channel Migration. *J. Marketing Res.* **45**(1) 60-76.

Barwise, P., J.U. Farley. 2005. The state of interactive marketing in seven countries: interactive marketing comes of age. *J. interactive Marketing*. **19**(3) 67-80.

Bass, F.M., N. Bruce, S. Majumdar, B.P.S. Murthi. 2007. Wearout effects of different advertising themes: a dynamic Bayesian model of the advertising-sales relationship. *Marketing Sci.* **26**(2) 179-195.

Bass, F.M., R.T. Lonsdale. 1966. An exploration of linear programming in media selection. *J. Marketing Res.* **3**(2) 179-188.

Bass, F.M., A. Krishnamurthy, A. Prasad, S.P. Sethi. 2005. Generic and brand advertising strategies in a dynamic duopoly. *Marketing Sci.* **24**(4) 556-568.

Bean, C.H. 1912. *The Curve of Forgetting*. Archives of Psychology, New York.

Berkowitz, D., A. Allaway, G. D'Souza. 2001a. The impact of differential effects on the allocation of advertising budgets across media. *J. Advertising Res.* **41**(2) 27-36.

Berkowitz, D., A. Allaway, G. D'Souza. 2001b. Estimating differential lag effects for multiple media across multiple stores. *J. Advertising.* **30**(4) 59-65.

Bhargava, M., N. Donthu. 1999. Sales response to outdoor advertising. *J. Advertising Res.* **39**(4) 7-18.

Brown, R.L., J. Durbin, J.M. Evans. 1975. Techniques of Testing the Constancy of Regression Relationships over Time. *J. Royal Statistical Society, Series B.* **37** 141-92.

Carroll, V.P., A.G. Rao, H.L. Lee, A. Shapiro, B.L. Bayus. 1985. The Navy enlistment marketing experiment. *Marketing Sci.* **4**(4) 352-374.

Chintagunta, P.K., N. Vilcassim. 1994. Marketing investment decisions in a dynamic duopoly: a model and empirical analysis. *International J. Res. in Marketing.* **11** 287-306.

Danaher, P.J., R.T. Rust. 1994. Determining the optimal level of media spending. *J. Advertising Res.* **34**(1) 28-34.

Davidson, R., J.G. MacKinnon. 2003. *Econometric Theory & Methods.* Oxford University Press, Oxford, UK.

De Fleur, M.L. 1956. A Mass Communication Model of Stimulus Response Relationships: An Experiment in Leaflet Message Diffusion. *Sociometry*. **19**(1) 12-25.

Dekimpe, M.G., D.M. Hanssens. 1995. The persistence of marketing effects on sales. *Marketing Sci.* **14**(1) 1-21.

Dertouzos, J.N., S. Garber. 2006. Effectiveness of advertising in different media. *J. Advertising.* **35**(2) 111-122.

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

Duffy, B.C. 1938. A look at advertising media. *J. Marketing.* **3**(2) 137-145.

Eastlack, J.O., A.G. Rao. 1986. Modeling response to advertising and pricing changes for "V8" cocktail vegetable juice. *Marketing Sci.* **5**(3) 245-259.

Ewing, M.T., E. du Plessis, C. Foster. 2001. Cinema advertising re-considered. *J. Advertising Res.* **41**(1) 78-85.

Fechner, G.T. 1851. Zend-Avesta oder über die Dinge des Himmels und des Jenseits, Voss, Leipzig. Vol. **2** 373–386.

Fechner, G.T. 1887. Über die psychischen Massprincipien und das Weber'sche Gesetz. *Philosophische Studien.* **4** 161–230.

Feichtinger, G., R.F. Hartl, S.P. Sethi. 1994. Dynamic optimal control models in advertising: recent developments. *Management Sci.* **40**(2) 195-226.

Gatignon, H., D.M. Hanssens. 1987. Modeling marketing interactions with application to salesforce effectiveness. *J. Marketing Res.* **24**(3) 247-257.

Gensch, D.H. 1970. Media factors: a review article. *J. Marketing Res.* **7**(2) 216-225.

Hanssens, D.M., H.A. Levien. 1983. An econometric study of recruitment marketing in the U.S. Navy. *Management Sci.* **29**(10) 1167-1184.

Harvey, A.C. 1994. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge, UK.

Hugues, M. 1975. An empirical study of media comparison. *J. Marketing Res.* **12**(2) 221-223.

Jagpal, H.S. 1981. Measuring joint advertising effects in multiproduct firms. *J. Advertising Res.* **21**(1) 65-69.

Jastram, R.W. 1955. A treatment of distributed lags in the theory of advertising expenditure. *J. Marketing*. **20**(1) 36-46.

Kamien, M.I., N.L. Schwartz. 1991. *Dynamic Optimization*. Elsevier Science Publishing Company, New York.

Krinsky, I., A.L. Robb. 1986. On approximating the statistical properties of elasticities. *Rev. Econ. and Statistics*. **68**(4) 715-720.

Leone, R.P. 1995. Generalizing what is known about temporal aggregation and advertising carryover. *Marketing Sci.* **14**(3, Part 2 of 2) G141-150.

Little, J.D.C., L.M. Lodish. 1969. A media planning calculus. *Operations Research*. **17**(1) 1-35.

Lodish, L.M., M. Abraham, S. Kalmenson, J. Livelsberger, B. Lubetkin, B. Richardson, M.E. Stevens. 1995. How TV advertising works: a meta-analysis of 189 real world split cable TV advertising experiments. *J. Markeketing Res.* **32**(2) 125-139.

Lucas, D.B., S.H. Britt. 1950. *Advertising Psychology and Research.* McGraw-Hill Book Co, New York.

Luo, X., N. Donthu. 2001. Benchmarking advertising efficiency. *J. Advertising Res*. **41**(6) 7-18.

Montgomery, D.B., A.J. Silk. 1972. Estimating dynamic effects of market communications expenditures. *Management Sci.* **18**(10) B485-B501.

MSI. 2002. *Research Priorities 2002-2004*. Marketing Science Institute, Cambridge, USA.

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

Naik, P.A., K. Raman. 2003. Understanding the impact of synergy in multimedia communications. *J. Marketing Res.* **40**(4) 375-388.

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

Naik, P.A., D.E. Schultz, S. Srinivasan. 2007. Perils of using OLS to estimate multimedia communication effects. *J. Advertising Res.* **47**(3) 257-269.

Nerlove, M., F.V. Waugh. 1961. Advertising without supply control: some implications of a study of the advertising of oranges. *J. Farm Economics.* **43**(4) 813-837.

Palda, K.S. 1965. *The Measurement of Cumulative Advertising Effect*. Prentice-Hall, Englewood Cliffs.

Raman, K., P.A. Naik. 2004. Long-term profit impact of integrated marketing communications program. *Rev. Marketing Sci.* **2**(8).

Scott, W.D. 1903. *The Psychology of Advertising*. Boston.

Sethi, S.P., T. Thompson. 2000. *Optimal Control Theory - Applications to Management Science and Economics*. 2nd ed. Kluwer Acadamic Publishers, Amsterdam, NL.

Smith, T.M., S. Gopalakrishna, R. Chatterjee. 2006. A three-stage model of integrated marketing communications at the marketing-sales interface. *J. of Marketing Res.* **43**(4) 564-579.

Srinivasan, S. 1976. Decomposition of a multi-period media scheduling model in terms of single period equivalents. *Management Sci.* **23**(4) 349-360.

Teng, J.-T., G.L. Thompson. 1985. Optimal strategies for general price-advertising models. Optimal Control Theory and Economic Analysis (G. Feichtinger, ed.). Vol. 2. Amsterdam: North-Holland.

Tull, D.S. 1965. The carry-over effect of advertising. *J. Marketing.* **29**(2) 46-53.

Vakratsas, D. 2005. Advertising response models with managerial impact: an agenda for the future. *Appl. Stoch. Models in Bus. & Ind.* **21** 351-361.

Vakratsas, D., F.M. Feinberg, F.M. Bass, G. Kalyanaram. 2004. The shape of advertising response functions revisited: a model of dynamic probabilistic thresholds. *Marketing Sci.* **23**(1) 109-119.

Vakratsas, D., Z. Ma. 2005. A look at the long-run effectiveness of multimedia advertising and its implications for budget allocation decisions. *J. of Advertising Res.* **45**(2) 241-254.

White, H. 1982. Maximum likelihood estimation of misspecified models. *Econometrica*. **50**(1) 1-25.