Marketing Science Institute Working Paper Series 2011
Report No. 11-100

The Complexity of Multi-media Effects

Ceren Kolsarici and Demetrios Vakratsas

Report Summary

The measurement of multi-media effects has become a top priority for academics and practitioners. This is a challenging task since response to advertising is a complex phenomenon characterized by thresholds, saturation, medium-specific effects, and media interactions. Research on the subject—as well as practical applications—have been limited in terms of the number of media included (usually up to three) and the types of effects allowed.

In this report, authors Kolsarici and Vakratsas offer the first comprehensive investigation of multi-media effects using a flexible and easily implemented statistical learning methodology: multivariate adaptive regression splines (MARS). Their empirical application employs both time-series and cross-sectional data for durable goods (cars) and nondurable packaged goods (beer)—totalling over 50 brands—as well as the Leading National Advertisers (LNA) database. All data sets include monthly information on unit sales and advertising expenditure for various media.

Their findings offer rich managerial implications and suggest that the standard nomenclature based on "concave," "convex," and "S-shaped" media effects is inadequate. Rather, multi-media effects are better described using "hockey-stick," "V-shaped," and "inverted V-shaped" functions and their combinations. Another substantive finding is that media interactions may have a detrimental effect on sales—suggesting that cross-media effects are not always *synergistic* but could be *antagonistic* as well. The authors call this phenomenon *interaction super-saturation*, and attribute it to simultaneous exposure to frequently used or saturated media. Thus, media synergies should not be considered as the de facto preferred outcome, nor should they be assumed a priori when considering the issue of optimal budget allocation.

A final important, and alarming, finding is that more than 50% of the total media investment is inefficient (below threshold, beyond saturation, or in declining segments of the response), unfortunately confirming the old adage "half of my advertising is wasted." These results underline the need for a comprehensive examination of multi-media effects by marketing managers. The findings point to potential reasons for inefficient allocation: (1) relying on aggregated (total) rather than medium-specific response, (2) ignoring detrimental interaction effects, and (3) using standard methodologies that cannot capture the complexity and diversity of multi-media effects.

The proposed methodology MARS offers three improvements: it performs well even for a large number of media, it allows for easy calculation of turning points (threshold, saturation, and super-saturation) and detection of inefficient investment allocation, and it can be easily implemented using available software. These advantages will hopefully spark more applications in marketing practice and encourage further academic research. An immediate priority of future research should be the study of optimal media allocation decisions when brands are faced with an array of response functions for different media, as shown to be the case in this study.

Ceren Kolsarici is Assistant Professor, Queen's University Business School, Kingston, Ontario, Canada and Demetrios Vakratsas is Associate Professor and Quebec Teaching Chair, Desautels Faculty of Management, Montreal, Quebec, Canada.

The Complexity of Multi-Media Effects

Introduction

Media proliferation and the increasing use of Integrated Marketing Communication (IMC) programs have made the measurement of multi-media effects a top priority for academics and practitioners ("Marketing Science Institute Research Priorities," 2008). However, comprehensive evaluation of multi-media effects presents researchers and marketing practitioners with the following challenges:

1) *Market response to advertising is a complex phenomenon.* Many research studies have shown that market response to advertising is frequently characterized by thresholds, saturation and super-saturation, requiring flexible response functions (e.g. Ackoff & Emshoff, 1975; Dubé, Hitsch, & Manchanda, 2005; Vakratsas, Feinberg, Bass, & Kalyanaram, 2004).

2) *Response may vary by medium*. For example, magazine advertising may exhibit high thresholds due to the high level of clutter in the medium, but late saturation due to its predominantly informative content (e.g. Dijkstra, Buijtels, & van Raaij, 2005; Vakratsas & Ma, 2005). On the other hand, cable TV advertising may exhibit no thresholds but early saturation due to potential duplication with network TV audiences who are exposed to the same advertising. Thus, response shape may be unique to each medium, requiring a different functional form to capture its effects on sales (e.g. Eastlack & Rao, 1986).

3) *Complexity of media interactions.* If marginal response to each medium is characterized by complexity, then the same should be expected of interactive effects. Yet, the few studies that have considered media interaction effects typically assume linear or

monotonic interactions (e.g. Naik & Raman, 2003). Although the implicit assumption in using monotonic interactions is that multi-media spending creates synergies, a competing argument is that it may also create negative effects since simultaneous exposure to multiple media can lead to faster saturation.

4) *High dimensionality*. The previous points suggest that measuring multi-media effects requires flexible multivariate models capable of taking on different shapes for each medium. However, flexible models in high dimensions (number of variables) usually suffer from the curse of dimensionality (Bellman, 1961). Thus, modeling of complex multi-media effects requires a sophisticated, yet parsimonious approach that avoids the curse of dimensionality.

In this study we address these challenges by launching a comprehensive investigation of complex multi-media effects, an issue of great importance to marketing managers. We use an easily implementable but flexible modeling approach, Multivariate Adaptive Regression Splines (MARS), due to Friedman (1991), to examine the effects of a large number of media, including those with relatively low levels of allocation. To date, industry studies on multi-media effectiveness have considered a limited number of media due to data constraints and the large number of possible interactions (e.g. Havlena, 2008). Hence, our work makes the following contributions:

a) It is the first to comprehensively examine potentially complex effects of a large number of media. Our empirical application uses data covering consumer durables (cars), packaged goods (beer) and the top corporate advertisers (Leading National Advertisers). Thus, we are able to provide reliable empirical evidence for phenomena such as multiple thresholds, early and super-saturation for multiple media in a variety of settings.

b) It proposes, based on the empirical findings, a new nomenclature of media response shapes featuring a repertoire of "hockey-stick," "V"-shapes and their combinations. This challenges the existing typology of "concave" "convex" and "S-shaped" responses (e.g. Simon & Arndt, 1980), which appears to be inadequate.

c) It provides critical managerial implications regarding media investment efficiencies, calculated using the shapes of media effects under the proposed nomenclature. Specifically, our findings suggest that more than 50% of media investment allocation is inefficient (below threshold, beyond saturation or in declining segments of the response). This suggests that measurement of multi-media effects has not received sufficient attention in practice, or it has not been implemented using appropriate methodologies.

The rest of the paper proceeds as follows. First, we briefly summarize the relevant literature with respect to multi-media effects and advertising response complexity. We then introduce the proposed methodology (MARS) and proceed with a description of the data and the model specification. We follow with the discussion of the main findings and conclude with the implications of our study and suggestions for future research.

Related Literature

Despite the proliferation of media alternatives in marketing practice, there is a relative paucity of research on multi-media effects. The majority of marketing science literature has mainly focused on the effects of total or single-medium advertising spending. Studies on single medium or total advertising spending agree on the complexity of its effects (Bemmaor, 1984; Hanssens, Parsons,

& Schultz, 2001; Simon & Arndt, 1980; Vakratsas, et al., 2004). For example, Vakratsas et al. (2004) provide empirical evidence that market response to advertising is not necessarily globally concave and advertising thresholds indeed exist, particularly for evolving product categories. Dubé, Hitsch, & Manchanda (2005) also offer support for the existence of threshold effects using a spline approach. Moreover, experimental studies on advertising response found evidence for positive effects of decreased advertising levels, leading to patterns including V-shaped response and bi-modal M-shaped response (e.g. Ackoff & Emshoff, 1975; Hahn, et al., 1992).

Research on multi-media advertising has mainly focused on the relationship between a few selected media and has investigated how the simultaneous use of multiple media affects marketing performance measured by sales, market share, or awareness. The concept of synergy, defined as the combined effect of two or more media exceeding the sum of their individual effects and operationalized through the interaction effect, is a focal point of interest in these studies. Regarding traditional media, research has shown positive interactions between TV and radio (Edell & Keller, 1989), TV and print (Confer & McGlathery, 1991; Naik & Raman, 2003), TV and direct mail (Stafford, Lippold, & Sherron, 2003), as well as radio and newspaper (Jagpal, 1981) advertising using mainly laboratory experiments. Naik and Raman (2003) examine the synergies between TV and print advertising and not only provide empirical evidence for the existence of cross-media synergies in multi-media communications, but also underline their implications for budget allocation decisions. A seemingly counterintuitive implication of Naik and Raman's study is that managers should decrease (increase) the proportion of media budget allocated to the more (less) effective communications strategy. This finding highlights the significance of considering media with a smaller contribution, a point also suggested by Eastlack and Rao (1989).

Recently, following the growth of new media, researchers have focused on the relationship between online and offline advertising (e.g. Dijkstra, Buijtels, & van Raaij, 2005; Havlena, Cardarelli, & Montigny, 2007; Naik & Peters, 2009; Stephen & Galak, 2009). Dijkstra et al. (2005), using laboratory settings, examine traditional and new media interactions by investigating complementary effects among TV, print and Internet advertising in creating cognitive, affective and conative responses. Their results suggest that TV-only and print-only campaigns are at least as effective as the multi-media ones, while internet advertising benefits from additional media support and is not superior to multi-media campaigns These mixed results point to the complexity of media interactions due to a variety of factors such as the type of media involved, amount of exposure, etc. A recent study by Naik and Peters (2009) uses a hierarchical model of multi-media communications in which within and cross-media interactions are considered between offline (TV, print, and radio) and online (search and banner) ads. The results reveal synergies between online and offline media, suggesting that the media budget should be increased with online spending taking the biggest share of the raise. Although this article is the first study to consider the effects of a large number of media, the complexity of such effects is not investigated and response functions are assumed to have straightforward patterns (linear, log-linear). Finally, Vakratsas and Ma (2005) examine the long-term effects of three media (magazine, network TV and spot TV) and their allocation implications for the top two brands in the SUV market but did not consider interaction effects due to sample size restrictions. They conclude that magazine and network advertising exhibit positive long-term effects (especially for the market leader) whereas spot advertising has a negative long-term effect.

Although the aforementioned research has advanced our understanding on multi-media effects, it has a considerable number of limitations. First, the reviewed studies, with the exception of Naik and Peters (2009), have focused on a limited number of media, which is far from a realistic representation of the current marketing environment. Second, the response functions considered are relatively simple (i.e. concave and linear) and the interaction effects, when acknowledged, are assumed to be monotonic. Finally, the possibility that media may produce diverse response shapes has been largely ignored. These are important issues to address in order to obtain a more comprehensive and accurate picture of the complexity of multi-media effects and their implications for marketing managers. The Multivariate Adaptive Regression Splines (MARS) methodology, briefly outlined in the next section, can help researchers and managers to effectively deal with these issues.

Methodology[1]

We propose the use of Multivariate Adaptive Regression Splines (MARS) to capture the complexity of multi-media effects. MARS is a flexible, non-parametric, statistical learning method due to Friedman (1991) and has been shown to offer substantial improvement over other commonly employed non-parametric methodologies, such as standard splines and Kernel regression, in moderate sample sizes (i.e. 50 to 1000) and moderate to high dimensions (3 to 20). A distinctive feature of MARS is its reliance on the arithmetic concepts of addition and multiplication in model building (Hastie, Tibshirani, & Friedman, 2001), which allows it to be more parsimonious and sidestep the "curse of dimensionality" issue. Specifically, MARS is an

[1] The interested reader may refer to Appendix 1 for the technical details.

adaptive algorithm, based on recursive partitioning, which dynamically adjusts its strategy to take into account the behavior of the function to be approximated. The response function is estimated by optimally dividing the domain for each predictor variable (i.e. each medium) and fitting univariate splines (basis functions[2]) at each sub-region as shown in Figure 1. The partitioning of each variable is conditional on the partitioning of all the other variables, thus guaranteeing optimality across all dimensions. Recursive partitioning is a powerful paradigm owing to its ability to exploit the low local dimensionality of functions. For instance, although a function may depend on a large number of variables globally, the number of these might decrease significantly for any given local region. Moreover, the regions become more and more local as the recursive splitting proceeds. (Figures and tables follow References.)

The MARS algorithm works very similarly to a forward stepwise linear regression followed by backward elimination to control for over-fitting. However, instead of the original variables, basis functions and their interactions are introduced into the model which has the following form:

$$\hat{f}(\mathbf{X}) = \beta_0 + \sum_{m=1}^{M} \beta_m l_m(\mathbf{X}) \qquad (1)$$

where X is the vector of media, l(X) is the basis function in the form of a univariate spline and M is the total number of basis functions in the model.

MARS has a number of key advantages in tackling high dimensional problems. First, the regression surface is built up parsimoniously, using the main effects and interactions locally to avoid bias-prone representations (Hastie et al. 2001). Second, MARS employs a hierarchical forward selection strategy, where higher order interactions are only built up from the terms

[2] For the definition of technical terms refer to the glossary at the end of the report.

already in the model. This unique characteristic of MARS not only increases the efficiency of the search by avoiding the exploration over an exponentially growing space of alternatives, but also greatly facilitates the model's interpretability, which is a common flaw for other nonparametric methods. Third, unlike Kernel regression, boundary effects and curse of dimensionality are not applicable to MARS since it operates based on arithmetic concepts rather than geometric ones. The aforementioned properties of MARS make it a front-runner among methods for modeling multi-media effectiveness. In the following section we discuss our empirical application and provide the exact model specification.

Empirical Application

Data

We investigate the complexity of multi-media effects using various data sets from the US market. In the interest of providing generalizable conclusions we consider both time-series and cross-sectional data concerning durables (cars), packaged goods (beer) – as well as using the well-known leading national advertisers (LNA) data base. Time-series data are drawn from two subcategories of automobiles: Sports Utility Vehicles (SUVs) and hybrids. In the SUV category, we focus on the two top selling brands, Ford Explorer and Jeep Grand Cherokee. Similarly, in the hybrid subcategory we focus on the top three car models, namely Toyota Prius, Honda Civic and Toyota Camry. Data are available since the introduction of each brand. All data sets include monthly information on unit sales and advertising expenditures for various media. The source for sales data is Ward's Automotive Yearbook, whereas TNS media intelligence is the source of advertising data.

We also use cross-sectional data sets, the first covering 44 beer brands in US market for the year 2001. The data are part of a larger data set made available by Information Resources Inc. (Bronnenberg, Kruger, & Mela, 2008). The beer category was selected due to the high number of active brands with a substantial amount of variation in advertising levels for a wide range of media. The data for each brand are pooled across stores and aggregated at the annual level. Matching annual advertising data were obtained by TNS Media Intelligence. The second cross-sectional data set consists of annual sales and media spending for the Leading National Advertisers in US for the year 2002 ("2003 Leading National Advertisers Report," 2003).

Table 1 provides summary information for all data sets with the shaded cells highlighting the media selected for the analysis. The selection process considered each medium's spending allocation as percentage of total spending and its frequency of use by examining the cost of unit advertising. Moreover, we excluded some media with low allocation that were highly correlated with included media, to facilitate interpretability (Friedman, 1991). Despite the elimination of certain media from the analysis we still have a high number of media, which allows us to thoroughly evaluate the performance of MARS and alternative benchmark methods. It should be noted that in the case of MARS the media included in the final model may be fewer than those considered due to its backward elimination process.

Model specification and estimation

The employed MARS specification has the following form:

$$S_i = \hat{f}(\mathbf{A}_i) = \beta_{0i} + \sum_{p=1}^{P}\sum_{m=1}^{M}\beta_{pmi}\prod_{k=1}^{K_m}\left[s_{p(k_m)i}\left(A_{p(k_m)i} - t_{p(k_m)i}\right)\right]_+ \qquad (2)$$

where bold letters represent matrices and:

S_i, is sales in units for observation i, i=1,2,…N;

A_p, is the advertising spending variable for medium p, $p \in \{1,2,...,P\}$;

M, is the number of basis functions in the final model;

K_m, is the number of splits that gave rise to the m^{th} basis function (i.e. level of interaction, s.t. $m \in \{1,2,...,M\}$)

t_{k_m}, is the knot point for the k^{th} interaction term of m^{th} basis function.

It should be noted that the right-most term in brackets corresponds to the basis functions (i.e. 1_m in Equation (1)), and the subscript i indexes time for time series data and brand/firm for cross sectional data sets.

In equation (2), β_0 is the coefficient of the constant basis function (intercept) and the sum is over the basis functions retained after the backward elimination process. The term s_{km} takes on values ± 1 to imply right or left of the associated reflected pair.

For the time series data, we use a goodwill (stock) advertising variable defined as follows:

$$G_{pi} = (1 - \phi_p) G_{p(i-1)} + A_{pi} \tag{3}$$

Where A_{pi} is defined as above and φ_p is the decay constant for medium p. The reason for using stock, instead of flow, variables for advertising spending in time-series data is two-fold. First, it enables us to account for the temporal effects of ad spending and, second, eliminate problems due to zeros. Based on past literature and preliminary analyses, we use φ=0.5 for magazines and φ =0.9 for all the other media. In the analysis of cross-sectional data sets, the advertising variable in Equation (2) is the actual spending amount, A. Also, following Friedman (1991), we choose the degrees of freedom, which represent the cost of each basis function optimization, to be between 2 to 4 for each MARS model. The model was estimated using the MARS 3.0 software (Salford Systems, 2008).

Findings

Main effects

The estimated main effects are summarized in Table 2, including response shape characterization, turning points and inefficient allocation.

The functions illustrated in Figure 2, panels (a) to (h), exemplify the diversity of the estimated responses across media and data sets. The X-axis represents the investment range for the corresponding medium, while the Y-axis represents contribution to sales. The dots mark threshold (T) and saturation (S) points. We define threshold as a local minimum, above which the incremental effect of a marginal increase in spending on sales is positive, and saturation as a local maximum below which the incremental effect of a marginal increase in spending on sales is positive. When the incremental effect of marginal spending increase *above* the saturation point is negative, then we have the case of super-saturation. One of the advantages of MARS is the ease of calculation of such turning points using the estimated knots. For instance, in Figure 2(a), which exhibits the Network TV effect for Honda Civic Hybrid, the threshold level corresponds to the knot point of the increasing basis function. The same principle works for more complex response shapes such as that of Figure 2(g). The turning points were confirmed with a more conventional approach of using the first derivatives of the estimated functions (see Appendix 1). We do not classify any boundary points as threshold or saturation as we cannot speculate what happens below (lowest) or above (highest) them. However, absence of threshold or saturation points does not preclude inefficient allocation. For example, the right hockey-stick function of Figure 2(b) does not have any threshold or saturation points, yet the entire allocation is inefficient since it produces a non-increasing response.

As seen in Table 2 and Figure 2, response shapes vary from relatively straightforward hockey-stick-types, to non-monotonic V-types, and ultimately more complex ones involving their combinations (e.g. panels (g) and (h)). The derived nomenclature of media effects shown in Table 2 and illustrated in Figure 2 is indicative of the diversity of media effects. Thus, standard typologies based on "convex," "concave," and "S-shaped" functions (e.g. Simon & Arndt, 1980) appear to be inadequate. Validation tests comparing MARS to commonly employed parametric and non-parametric methods offer further support for this claim (see Appendix 2 for a description of the benchmark models and Appendix 3 for extensive comparisons).

In addition to the diversity of multi-media effects already reflected in the proposed nomenclature, a few important points emerge from the examination of Table 2 and Figure 2. First, media response shapes are frequently non-monotonic, showing super-saturation and multiple turning points (e.g. Figures 2(e)-(h)). Super-saturation suggests that excessive efforts of advertisers turn consumers away (e.g Hanssens, et al., 2001). The bimodal patterns in Figures 2(g) and (h) may be attributed to the existence of multiple segments of consumers with distinct thresholds and saturation points (Ackoff & Emshoff, 1975). It is interesting to note that for the cases where a threshold is preceded by a saturation point, as in Figure 2(g), the range between the saturation and threshold points is both a super-saturation and a sub-threshold range, exemplifying the complexity of effects. This applies, for example, to cases involving a "V-shaped" response for the left side of the "V" (corresponding to the decreasing part of the response).

Second, "inverted hockey-stick"-type of responses, corresponding to early saturation or non-increasing effects, are quite common and indicate the presence of significant inefficient allocation. For example, the two cases of newspaper advertising have the lowest saturation levels, at around 5% of the maximum investment, suggesting that 95% of spending allocated to this medium is inefficient. This may explain the latest advertiser tendencies to decrease newspaper spending ("The Recession in Advertising," 2009).

Third, magazines feature prominently as an influential medium in all data sets with the exception of beer where they are sparingly used. This could be due to the complementary nature of magazine advertising which allows for potentially unlimited exposure time due to its longer life and the possibility of multiple exposures via the pass along rate (Katz, 2003). It should also be noted that they exhibit low or no saturation.

Finally, and perhaps most importantly from a managerial perspective, turning point (threshold and saturation) calculation, rounded to the nearest multiple of 5%, reveals that the majority (55%) of the total media investment allocation is inefficient (below threshold, beyond saturation or in response downslopes). This emphasizes the need for a comprehensive evaluation of multimedia investment which will lead to more efficient allocation. Threshold levels over all media and data sets range from 5% to 30%, while saturation levels have an average of approximately 50% excluding newspaper ads, which, as mentioned before, exhibit even higher saturation. Interestingly, spot TV advertising appears to have the most efficient allocation. This could be attributed to its "pick and choose" nature (e.g. Katz 2003) that allows advertisers to avoid saturated markets and audiences.

Interaction effects

Figure 3 illustrates two cases that typify the complexity of interactive effects, showing how the combined effect of two media can depart considerably from each medium's main (or marginal) effect. The first interaction (Figure 3(a)) involves two highly saturated media (cable TV and newspapers- see also Table 2). While for moderate levels of both media allocations the interaction effect is increasing (middle of top surface in Figure 3(a)), for high allocation levels of both media ("far corner" of surface) the effect decreases suggesting that the saturation exhibited by each medium has a detrimental effect on the interaction. We call this effect *interaction super-saturation*. However, even efficient media such as spot TV in figure 3(b) are susceptible to this effect. Although spot TV has an increasing marginal effect on sales response (Table 2), when combined with the highly saturated newspaper medium, it can produce decreasing response. In this case, one medium's (spot TV) marginal effect can be considerably altered through the interaction of another medium (newspapers), with the resulting combined effect not directly deriving from each medium's marginal effect. Hence, linear interactions would inadequately capture the combined media effects as model comparisons demonstrate in Appendix 3.

Thus, it appears that interactive effects are not necessarily *synergistic* but could be *antagonistic* as well due to interaction super-saturation, and considering only each medium's marginal effect provides a partial view of multi-media effectiveness. Thus, even if a medium may work well by itself, when its exposure is combined with another medium at excessive levels the effect could be negative. In sum, we find that interaction effects are complex and challenging to model using standard methodologies. The advantage of MARS is that it simplifies the modeling of complex effects by: (1) fitting combinations of simple linear functions to the observation domain and (2) eliminating the non-significant effects all together, hence increasing the interpretability of the results. Our validation tests in Appendix 3 provide formal support for these claims.

Implications and Conclusion

Substantive implications

Earlier empirical research (Eastlack and Rao 1986, 1989) suggested that response is likely to vary by medium and that future research should evaluate the effects of a larger number of media. Our work picks up on these suggestions and finds that response to media exhibits both diversity and complexity. In other words, sales response to media takes on different shapes, exhibits various turning points such as thresholds and saturation, and is frequently non-monotonic. Thus, the traditional nomenclature of "concave," "convex" and "S-Shaped" response is inadequate, since we frequently observe, "hockey-stick," "V-shaped," "inverted-V shaped," functions and their combinations. Furthermore, it is unlikely that response to total media allocation is representative of response to any particular medium. This is an important implication since inefficient investment ranges should be avoided for every medium and an aggregate (over all media) response function may not be able to provide such guidance.

Our implications on media interactions are also intriguing. Most prominently, we find that media interaction effects are not necessarily linear or monotonic as it is typically assumed in the literature. Non-monotonic interactions suggest that cross-media effects are not always positive. In other words, media interactions are not necessarily *synergistic* but could be *antagonistic* as well. We attribute this to the faster saturation triggered by simultaneous exposure to multiple media, which we call *interaction super-saturation.* Thus, media synergies should not be considered as the *de facto* outcome, nor should they be assumed *a priori* in models used for optimal allocation.

The fact that the majority (55%) of total media investment was found to be inefficiently allocated (below thresholds, beyond saturation and in downslopes) suggests that managers have not dealt effectively with this issue. Thus, unfortunately, the old adage “half of my advertising is wasted,” still applies. Our findings point to potential reasons for inefficient allocation: a) relying on aggregated (total) rather than medium-specific response, which would point to the sources of inefficiency, b) ignoring detrimental interaction effects such as interaction super-saturation, and c) using standard methodologies that cannot adequately capture the complexity of multi-media effects. Hopefully, the proposed MARS methodology can serve as a managerial blueprint on how to deal with the complexity of multi-media effects. A practical advantage of MARS is that turning points and, consequently, inefficient allocation can be easily calculated using the estimated knots. Moreover, MARS can be easily implemented using commercially available software, thus offering a “simple approach to complexity.”

Methodological implications

A major hurdle in evaluating multi-media effectiveness has been the issue of high dimensionality, which requires the implementation of flexible yet parsimonious methodologies. MARS, advocated in this study, satisfies both criteria of flexibility and parsimony. Our extensive validation tests confirm this as they clearly show that: (1) MARS provides a better fit and shows better average forecast performance than both parametric benchmarks and Kernel regression in all data sets and (2) the improvement of MARS over Kernel significantly increases as the number of media considered increases. Thus, our study offers a way forward for the systematic examination of multi-media effects, since the proposed model alleviates concerns regarding data density in a high number of dimensions.

In conclusion, while there is little doubt that multi-media effects are complex, their modeling has been a challenging task. This may well explain the relative dearth of research on the subject. In this study, responding to a call from MSI for more research on this issue, we propose a way of dealing with the challenges of modeling multi-media effects. The advocated methodology, MARS, performs well even for a large number of media, allows for easy calculation of turning points and detection of inefficient investment allocation, all advantages that make it a powerful managerial tool. This will hopefully spark more applications in marketing practice and encourage further academic research. An immediate priority of future research should be the study of optimal media allocation decisions when brands are faced with an array of response functions for different media, as shown to be the case in this study.

Appendices

Appendix 1 – MARS

MARS algorithm. The MARS algorithm works as follows. First, a set of piecewise linear basis functions, each of the form $(x-t)_+$ and $(t-x)_+$ where the subscript "+" refers to the positive part (i.e. $(x-t)_+ = \max(x-t,0)$), is created by forming reflected pairs for each variable X_p ($p \le P$) with knots at each observed value of that variable. Figure 1 depicts the basis functions $(x-0.5)_+$ and $(0.5-x)_+$, where the consecutive elements of the reflected pair are represented by the solid and the dashed lines respectively and the knot is located at t=0.5. The collection set, which includes all candidate functions to be added to the final model, can be represented as in Equation (A1) where p is the number of variables (i.e. dimensions) and N is the number of observations for each variable.

$$C = \left\{ (\mathbf{X}_p - t)_+ , (t - \mathbf{X}_p)_+ \right\}_{\substack{t \in \{x_{1p}, x_{2p}, \ldots, x_{Np}\} \\ p=1,2,\ldots,P}} \quad \text{(A1)}$$

The rest of the MARS algorithm works very similarly to a forward stepwise linear regression; however, instead of the original variables, basis functions, such as the ones included in the collection set C, and their interactions are introduced into the model, which has the following form:

$$\hat{f}(\mathbf{X}) = \beta_0 + \sum_{m=1}^{M} \beta_m l_m(\mathbf{X}) \quad \text{(A2)}$$

Starting with a constant function, $l_0(X) = 1$, a new basis function pair that produces the largest decrease in training error at each stage is introduced in the model. The products of all basis functions already in the model, $l_m(X)$, with each of the reflected pairs in the collection set, C, are considered as candidates for entry to the model at each selection step. Hence, the basis functions produced by the forward selection algorithm have the following form:

$$\hat{\beta}_{M+1} l_m(\mathbf{X})(\mathbf{X}_p - t)_+ + \hat{\beta}_{M+2} l_m(\mathbf{X})(t - \mathbf{X}_p)_+ \quad \text{(A3)}$$

The coefficients $\hat{\beta}_{M+1}$ and $\hat{\beta}_{M+2}$, are estimated by least squares and the process continues until the model reaches a pre-set maximum number of terms. The internal nodes of the tree represent different partitions, and the terminal nodes represent the final basis functions. The corresponding variable and the knot point for each partition are listed below the internal nodes. As seen in the binary tree representation, each final basis function is the product of the basis functions encountered in a traversal of the tree starting at the root of the tree and ending at its corresponding terminal node. In other words, in a MARS model higher order interactions are formed only for the terms that are already in the model, which significantly increases the efficiency of the procedure.

Following the forward selection procedure, a backward elimination process is applied to control for overfitting of Equation (A2) to the data. The term whose removal results in the smallest increase in the generalized cross-validation criterion (GCV) is deleted from the model at each stage. The GCV (Equation (A4)) is used as a lack-of-fit measure based on which the knot locations and the optimal number of basis functions, λ, are determined at the end of the backward elimination procedure:

$$\mathrm{GCV} = \frac{1}{N}\sum_{i=1}^{N} \frac{\left[y_i - \hat{f}_\lambda(\mathbf{X}_i) \right]^2}{\left(1 - \frac{C(\lambda)}{N} \right)^2} \quad \text{(A4)}$$

The numerator in Equation (A4) is the averaged-squared error while the denominator is a penalty function adjusting for the increased variance associated with higher model complexity.

C(λ) represents the effective number of parameters which, in addition to the number of terms in the model, takes into account the number of parameters used to determine optimal knot positions. Simulation studies have shown that one would pay the price of three parameters for selecting a knot in a piecewise linear regression. Hence, $\mathrm{C}(\lambda) = \mathrm{r} + \mathrm{cK}$, where *r* is the number of basis functions, *K* is the number of knots selected in the forward process, and c is equal to 3 (Hastie, et al., 2001). For further details on the MARS algorithm, the interested reader may refer to the original article (Friedman, 1991).

Calculation of turning points using first derivatives. We apply the MARS algorithm by constraining the interaction effects to be zero, and compute the threshold and saturation levels from the resulting MARS model after the backward eliminate. Therefore, the final MARS model consists, only, of the significant single variable basis functions.

Since MARS employs simple first degree polynomials to form the basis functions, we use the first derivatives of the corresponding main effect curves in our calculations. More specifically, we represent the threshold and saturation levels of each medium by the following mathematical formulations, where ε is a very small, positive real number and Y is the response measure (i.e. sales in units, in this chapter). Equations (A5) and (A6) represent the threshold - x_i^* - and saturation - x_i^{**} - levels for medium *i*, respectively:

$$x_i^* = \left\{ x_i \mid \frac{\partial Y}{\partial x_i}\bigg|_{x_i = x_i^* - \varepsilon} \le 0, \frac{\partial Y}{\partial x_i}\bigg|_{x_i = x_i^* + \varepsilon} > 0, \varepsilon > 0 \right\} \qquad \text{(A5)}$$

$$x_i^{**} = \left\{ x_i \mid \frac{\partial Y}{\partial x_i}\bigg|_{x_i = x_i^{**} - \varepsilon} > 0, \frac{\partial Y}{\partial x_i}\bigg|_{x_i = x_i^{**} + \varepsilon} \le 0, \varepsilon > 0 \right\} \qquad \text{(A6)}$$

In this operationalization, a threshold point is the local minimum point above which the incremental effect of a marginal increase in spending on sales is positive. Following a similar logic, saturation the local maximum point below which the incremental effect of a marginal increase in spending on sales is positive.

Appendix 2 – benchmark methods

Benchmark parametric models. Parametric models are frequently used to study advertising effects (Leeflang, Wittink, Wedel, & Naert, 2000). Although inflexible, they have several desirable properties such as consistency and asymptotic efficiency (Leeflang, et al., 2000, p. 397), in addition to their ease of interpretation and requirement of relatively few data points, when the true underlying function is close to the pre-specified parametric one and the number of dimensions is low. However, their specification is frequently laden with uncertainty, resulting in biased and inconsistent estimates.

One parametric benchmark we use is the semi-log model of advertising goodwill, which allows for decreasing returns to scale:

$$S_i = \alpha_0 + \sum_{p=1}^{P} \alpha_{1p} \ln(A_{pi}) \tag{A7}$$

for i=1,…,N and p=1,…,P.

The other parametric benchmark is the multiplicative model represented in Equation (A8). Multiplicative sales models have been quite popular in empirical marketing research, mainly due to their ability to accommodate various response shapes based on the value of the estimated coefficient (i.e. increasing returns to scale if $\delta>1$, and decreasing returns to scale if $0<\delta<1$ in Equation (A8)), and higher order interactions.

$$S_i = \exp(\delta_0)\prod_{p=1}^{P}(A_{pi})^{\delta_{1p}} \qquad \text{(A8)}$$

We estimate Equations (A7) and (A8) by ordinary least squares, which is a robust method with respect to predictive fit for these classes of models.

Benchmark non-parametric model. As a non-parametric benchmark we use the Kernel method with Gaussian kernels (Abe, 1995; Van Heerde, Leeflang, & Wittink, 2001), and its multivariate extension, the product kernel, proposed by Hardle (1999). The regression model can then be represented by the following formulation using the Nadaraya-Watson estimator (Nadaraya, 1970; Watson, 1964).

$$S^* = \frac{\sum_{i=1}^{N} S_i K\left(\frac{\mathbf{A}_i - \mathbf{A}^*}{\mathbf{h}}\right)}{\sum_{i=1}^{N} K\left(\frac{\mathbf{A}_i - \mathbf{A}^*}{\mathbf{h}}\right)} \qquad \text{(A9)}$$

In Equation (A9), the kernel function is,

$$K\left(\frac{\mathbf{A}_i - \mathbf{A}^*}{\mathbf{h}}\right) = K(A_{1i}, A_{2i}, \ldots, A_{Pi}) = (2\pi)^{-\frac{1}{2}P}\prod_{p=1}^{P} e^{-\frac{1}{2}\left(\frac{A_{pi}-A_p^*}{\mathbf{h}}\right)^2} \qquad \text{(A10)}$$

Consistent with previous notation, the bold letters represent vectors and:

$\mathbf{A}^*$**,** is the vector of media spending variables (i.e. goodwill for the time series data, and nominal spending for the cross-sectional data) for which the response is estimated;

$\mathbf{A}_i$**,** is the vector of media spending variables for which the response is observed;

p is the number of media in the model (i.e. dimensions);

h is the vector representing the bandwidth (i.e. smoothing constant) for all media variables[3].

Appendix 3 – Model Fit and Validation

In-sample performance. We report mean-squared errors (MSEs) for all models in Table A1 [4]. The first column displays MARS MSEs and the columns on the right report its percentage difference from benchmark models. These results indicate that non-parametric estimation improves fit to the sample data compared to both parametric models. This is largely expected due to the greater flexibility of non-parametric methods in capturing complex and/or non-monotonic media effects.

More interesting observations, however, arise from the comparisons between the two non-parametric methods. MARS achieves an average improvement of around 55% in mean squared errors over Kernel. The differences are highest for Toyota Camry hybrid and Leading National Advertisers, the former being the "shortest" (smallest number of observations), and the latter being the "widest" (highest number of variables), of the data sets analyzed with 33 observations and 7 media respectively. The superiority of MARS should be attributed to its arithmetic-based mechanism instead of the geometric-based approach of Kernel. Wide and short data sets represent the most challenging cases in terms of the available statistical power. Hence, the markedly better performance of MARS highlights its ability to deal with the *curse of dimensionality* and scarcity of data in general.

---

[3] The normal reference rule due to Silverman (1986) is used for bandwidth selection. Given our assumption of Gaussian Kernel function, the optimal bandwidth is calculated as: $h_{opt} = \left\{ \frac{4}{n(2p+1)} \right\}^{1/(p+4)}$

[4] We also used Mean Absolute Deviation (MAD) and Mean Absolute Percentage Deviation (MAPD). The results were similar and are omitted for ease of exposition. The formulas for the fit statistics are:

$$\text{MAD} = \frac{1}{N}\sum_{n=1}^{N}\left|y_n - \hat{y}_n\right|,\ \text{MSE} = \frac{1}{N}\sum_{n=1}^{N}\left[y_n - \hat{y}_n\right]^2,\ \text{MAPD} = \frac{1}{N}\sum_{n=1}^{N}\frac{100\left|y_n - \hat{y}_n\right|}{y_n}$$

To further illustrate the handling of higher dimensions by MARS, we selected two data sets and calculated in-sample prediction values by adding one variable at a time. We then compared the marginal improvement of MARS over Kernel for each data set by sequentially increasing the dimensionality. The results are presented in Table A2 for Ford Explorer and LNA, respectively. For each data set, we added variables based on their allocated spending percentage, as shown in Table 1, starting with the biggest-spending media first.

It is clear that as the number of variables considered in each data set increases, the marginal improvement of MARS over Kernel increases as well. In fact, both methods perform notably close to each other for up to three dimensions, with the Kernel method achieving smaller MAD and MAPD values for the LNA data set. However, beyond three dimensions MARS performs notably better than Kernel, even though in the case of the LNA data we consider media beyond the number selected in the final MARS specification (up to six considered versus four selected). This further illustrates the ability of MARS to handle high-dimensional problems.

Hold-out Performance. Improvement in fit is virtually guaranteed with more flexible methods; therefore, a more stringent test involves performance comparison in validation samples. There is no exact rule for choosing the number of observations for the training set, used to fit the models, and the validation set, used to evaluate predictive performance. We use a 50-50 split which is considered typical (Hastie, et al., 2001; Van Heerde, et al., 2001).

Table A3 presents prediction errors for all models. For all data sets, MARS achieves better predictive validity than the parametric benchmarks. MSE values for Kernel regression, on the other hand, are larger than the best fitting parametric model for three data sets. Semi-log model attains smaller out of sample MSE than Kernel for Honda Civic Hybrid, and Ford Explorer. For the LNA data, Kernel's predictive fit is inferior to both semi-log and multiplicative models. Thus, the well-known bias-variance tradeoff in which the Kernel method is locked seems to impede its relative performance.

MARS achieves better out of sample fit than Kernel in all data sets. The average fit improvement is around 30% but varies across data sets. Naturally, intrinsic characteristics of the data sets, such as the range of each variable in the validation sample, affect marginal performance improvement by changing the approximate optimal bandwidth in Kernel.

In sum, the validation exercise confirms the superiority of MARS, which should be attributed to its ability to remedy the curse of dimensionality problem due to the way it builds up response surfaces.

Table A1: Fit Statistics for MARS and Benchmark Models

		MSE	Comparison to Benchmarks[5]		
		MARS	Kernel	Semi-Log	Multiplicative
	Estimation				
	Hybrid Cars				
TIME SERIES	H.Civic	3.05E+05	-58.1[6]	-79.7	-78.9
	T. Camry	3.90E+05	-74.7	-79.5	-85.0
	T. Prius	4.25E+06	-57.4	-58.9	-84.3
	Average		-63.4	-72.7	-82.7
	SUVs				
	Explorer	3.42E+06	-55.6	-78.6	-82.5
	Jeep	8.60E+06	-34.7	-56.8	-61.8
	Average		-45.2	-67.7	-72.2
CROSS SECTIONAL	*L. Nat. Adv.*	1.16E+08	-84.3	-93.1	-92.8
	Beer	1.05E+11	-42.9	-95.3	-98.9
	Overall Average		**-58.3**	**-77.4**	**-83.5**

[5] Negative numbers denote improvement by MARS.

[6] It should read: MARS MSE is 58.1% less than Kernel's

Table A2: Comparison of MARS and Kernel Performance in Various Dimensions

Ford Explorer

		Difference		Difference		Difference
	MARS	Kernel	MARS	Kernel	MARS	Kernel
	3-Dim		4-Dim		5-Dim	
MAD	2278	-5.35	1674.9	-21.44	1397	-33.44
MAPD	8.13	-8.65	6.5	-17.72	5	-36.38
MSE	8.30E+06	-11.70	5.14E+06	-35.26	3.42E+06	-55.52

Leading National Advertisers

		Difference		Difference		Difference		Difference
	MARS	Kernel	MARS	Kernel	MARS	Kernel	MARS	Kernel
	3-Dim		4-Dim		5-Dim		6-Dim	
MAD	17732	5.4	13712	-12.15	11881	-11.80	9812	-22.47
MAPD	120.1	1.1	89.7	-12.17	70.4	-18.13	64	-19.73
MSE	6.11E+08	-28.12	3.61E+08	-54.19	2.62E+08	-63.91	1.56E+08	-78.57

Table A3: Predictive Validity Statistics For MARS And Benchmark Models

		MSE	Comparison to Benchmarks		
		MARS	Kernel	Semi-Log	Multiplicative
	Validation				
TIME SERIES	*Hybrid Cars*				
	H.Civic	1.39E+06	-35.6	-12.6	-40.3
	T. Camry	2.10E+06	-8.7	-63.8	-69.1
	T. Prius	2.60E+07	-30.1	-84.5	-100.0
	Average		-24.8	-53.6	-69.8
	SUVs				
	Explorer	5.64E+07	-42.7	-11.9	-99.6
	Jeep	3.05E+07	-15.3	-12.9	-93.9
	Average		-29.0	-12.4	-96.8
CROSS SECTIONAL	*L. Nat. Adv.*	2.50E+09	-23.1	-8.8	-13.8
	Beer	2.80E+11	-67.1	-78.6	-99.9
	Overall Average		**-31.8**	**-39.1**	**-73.8**

REFERENCES

Marketing Science Institute Research Priorities 2008-2010. (2008). Retrieved December 12, 2009, from http://www.msi.org/research/index.cfm?id=43

2003 Leading National Advertisers Report. (2003, June 23). Retrieved December 12, 2009, from http://adage.com/abstract.php?article_id=37727

Abe, M. (1995). A Nonparametric Density Estimation Method for Brand Choice Using Scanner Data. *Marketing Science, 14*(3), 300-325.

Ackoff, R. L., & Emshoff, J. R. (1975). Advertising Research at Anheuser-Busch, Inc. (1963-68). *Sloan Management Review*(Winter), 1-15.

Bellman, R. E. (1961). *Adaptive Control Processes*. Princeton, NJ: Princeton University Press.

Bemmaor, A. C. (1984). Testing Alternative Econometric-Models on the Existence of Advertising Threshold Effect. *Journal of Marketing Research, 21*(3), 298-308.

Bronnenberg, B. J., Kruger, M. W., & Mela, C. (2008). The IRI Marketing Data Set. *Marketing Science, 27*(4), 745-748.

Confer, M. G., & McGlathery, D. (1991). The Research Study: The Advertising Impact Of Magazines In Conjunction With Television. [Article]. *Journal of Advertising Research, 31*(1), 64-64.

Dijkstra, M., Buijtels, H. E. J. J. M., & van Raaij, W. F. (2005). Separate and Joint Effects of Medium Type on Consumer Responses: A Comparison of Television, Print, and the Internet. *Journal of Business Research, 58*(3), 377-386.

Dubé, J.-P., Hitsch, G. J., & Manchanda, P. (2005). An Empirical Model of Advertising Dynamics. *Quantitative Marketing & Economics, 3*(2), 107-144.

Eastlack, J. O., & Rao, A. G. (1986). Modeling Response to Advertising and Pricing Changes for V-8 Coctail Vegetable Juice. *Marketing Science, 5*(3), 245.

Eastlack, J. O., & Rao, A. G. (1989). Advertising Experiments at the Campbell Soup Company. *Marketing Science, 8*(1), 57.

Edell, J. A., & Keller, K. L. (1989). The Information Processing of Coordinated Media Campaigns. *Journal of Marketing Research , 26*(2), 149-163.

Friedman, J. (1991). Multivariate Adaptive Regression Splines. *Annals of Statistics, 19*(1), 1-67.

Hahn, M., Park, W. C., & Macinnis, D. J. (1992). The Adaptive Information Processing Hypothesis: Accounting for the V-Shaped Advertisign Response Function. *Journal of Advertising, 21*(2), 37-46.

Hanssens, D. M., Parsons, L. J., & Schultz, R. L. (2001). *Market Response Models, Econometric and Time Series Analysis* (2 ed.). Boston, MA: Kluwer Academic Press.

Hardle, W. (1999). *Applied Non-Parametric Regression* (Vol. 19). Cambridge, UK: Cambridge University Press.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. New York: Springer Science and Business Media.

Havlena, W. J. (2008). *Measuring Marketing Effectiveness in a New Media World.* Paper presented at the MSI Conference, IESE Business School, Barcelona, Spain.

Havlena, W., Robert Cardarelli, & Montigny, M. D. (2007). Quantifying The Isolated And Synergistic Effects Of Exposure Frequency For TV, Print, And Internet Advertising. *Journal of Advertising Research, 47*(3), 215-221.

Jagpal, H. S. (1981). Measuring Joint Advertising Effects in Multiproduct Firms. [Article]. *Journal of Advertising Research, 21*(1), 65.

Katz, H. (2003). *The Media Handbook: A Complete Guide to Advertising Media Selection, Planning, Research, and Buying* Mahwah, NJ: Lawrence Erlbaum Associates.

Leeflang, P. S. H., Wittink, D. R., Wedel, M., & Naert, P. A. (2000). *Building Models for Marketing Decisions*. Boston: Kluwer Academic Publishers.

Nadaraya, E. A. (1970). Remarks on Non-Parametric Estimates for Density Functions and Regression Curves. *Theory of Probability and its Applications, 15*, 134-137.

Naik, P. A., & Peters, K. (2009). A Hierarchical Marketing Communications Model of Online and Offline Media Synergies. [doi: DOI: 10.1016/j.intmar.2009.07.005]. *Journal of Interactive Marketing, 23*(4), 288-299.

Naik, P. A., & Raman, K. (2003). Understanding the Impact of Synergy in Multimedia Communications. *Journal Of Marketing Research, 40*(4), 375-388.

The Recession in Advertising. (July 30th, 2009). *The Economist*.

Silverman, B. W. (1986). Density Estimation for Statistics and Data Analysis *Monographs on Statistics and Applied Probability* (Vol. 26). London: Chapman & Hall.

Simon, J. L., & Arndt, J. (1980). The Shape of the Advertising Response Function. *Journal of Advertising Research, 20*(4), 11-28.

Salford Systems (2008). MARS for Windows, Version 3.0.

Stafford, M. R., Lippold, E. M., & Sherron, C. T. (2003). The Contribution of Direct Mail Advertising to Average Weekly Unit Sales. *Journal of Advertising Research, 43*(2), 173-179.

Stephen, A., & Galak, J. (2009). The Complementary Roles of Traditional and Social Media Publicity in Driving Marketing Performance. *Insead Working Paper*

Vakratsas, D., Feinberg, F. M., Bass, F. M., & Kalyanaram, G. (2004). The Shape of Advertising Response Functions Revisited: A Model of Dynamic Probabilistic Thresholds. *Marketing Science, 23*(1), 109-119.

Vakratsas, D., & Ma, Z. (2005). A Look at the Long-run Effectiveness of Multimedia Advertising and Its Implications for Budget Allocation Decisions. *Journal of Advertising Research, 45*(2), 241-254.

Van Heerde, H. J., Leeflang, P. S. H., & Wittink, D. R. (2001). Semiparametric Analysis to Estimate the Deal Effect Curve. *Journal of Marketing Research, 38*(2), 197-215.

Watson, G. S. (1964). Smooth Regression Analysis. *Sankhya, Series A*(26), 359-372.

Glossary of Terms

Backward Elimination: The procedure used for the elimination of suboptimal basis functions. It starts with the full model and then removes basis functions one at a time based on the lowest error increase criterion.

Basis function: A local function, usually in the form of an n-degree polynomial, defined on a specific sub-region of the data domain (i.e. n=1 for linear; n=2 for quadratic; n=3 for cubic basis function etc.).

Forward Stepwise Regression: The procedure by which the MARS model is sequentially building up, starting with the intercept and adding the basis function that most improves the fit.

Knot: A point in the real line that divides the data domain into sub-regions.

Spline function: A function used for approximation, which is composed of segments of simple functions defined on sub-regions, and joined at their endpoints with a suitable degree of smoothness.

TABLES AND FIGURES

Table 1: Summary Description of Data

		Time Range	# of Observations	Magazine	Newspaper	Network TV	Spot TV	Syndicated TV	Cable TV	Radio	Internet	# of Media Included (Dimensions)
				MEDIA (% Spending)								
CROSS SECTIONAL	Lead National Adv.s.	2002	96	17.3	13.8	30.7	15.2	3.2	14.1	2.2	3.2	7
	Beer	2001	44	5.2	0.2	64.1	6.8	2.2	19.9	0.6	0.8	3
TIME SERIES	Ford Explorer	04/90-12/00	129	28.5	4.5	41.3	18.6	1.8	3.8	0.2	1.2	5
	Jeep Grand Cherokee	01/92-12/06	180	22.7	6.6	25.2	37.4	0.4	7	0.4	0.2	4
	Honda Civic Hybrid	04/02-12/08	81	18.7	1	34.7	35.3	0	9.4	0	0.8	4
	Toyota Camry Hybrid	02/06-12/08	35	34.1	0.2	49.9	5.1	0	7.8	0	2.9	3
	Toyota Prius Hybrid	01/01-12/08	96	32.2	7.7	27.1	19.5	0	6.3	0	7	3

Table 2: Summary of Main Effects

	Response Shape	Turning points: Threshold Level	Turning points: Saturation Level[7]	Inefficient Allocation
Magazine				
Toyota Prius Hybrid	V	10%	-	10%
Toyota Camry Hybrid	Inverted Right Hockey-stick	-	45%	55%
Honda Civic Hybrid	Left Hockey-stick-inverted-V	35%	65%	70%
Ford Explorer	Reverse Z	15%	40%	75%
Jeep Grand Cherokee	Left Hockey-stick-V combination	T1:30% T2:65%	40%	55%
Leading National Adv.	Kinked increasing with threshold	30%	-	30%
			Mean	**49%**
Network TV				
Toyota Prius Hybrid	Inverted Right Hockey-stick	-	15%	85%
Honda Civic Hybrid	Left Hockey-stick	30%	-	30%
Ford Explorer	Inverted left hockey-stick-Inverted-V	20%	55%	65%
Beer	Left Hockey-stick	15%	-	15%
			Mean	**49%**
Cable TV				
Ford Explorer	Left Hockey-stick-V combination	T1: 5% T2: 25%	20%	10%
Jeep Grand Cherokee	Right Hockey-stick	-	-	100%
Leading National Adv.	Left Hockey-stick-inverted-V	15%	25%	90%
Beer	Inverted-V	-	35%	65%
			Mean	**66%**
Spot TV				
Ford Explorer	Kinked increasing	-	-	0%
Jeep Grand Cherokee	W	T1: 5% T2:55%	20%	40%
Leading National Adv.	Left Hockey-stick	10%	-	10%
			Mean	**17%**
Newspaper				
Ford Explorer	Inverted Right Hockey-stick	-	5%	95%
Leading National Adv.	Inverted Right Hockey-stick	-	5%	95%
			Mean	**95%**
Internet				
Honda Civic Hybrid	Inverted Left Hockey-stick	-	-	100%
			Mean	**100%**
			Overall Mean	**55%**

[7] Includes super-saturation points as well.

Figure 1: Reflected Pair of Piecewise Linear Basis Functions Used by MARS

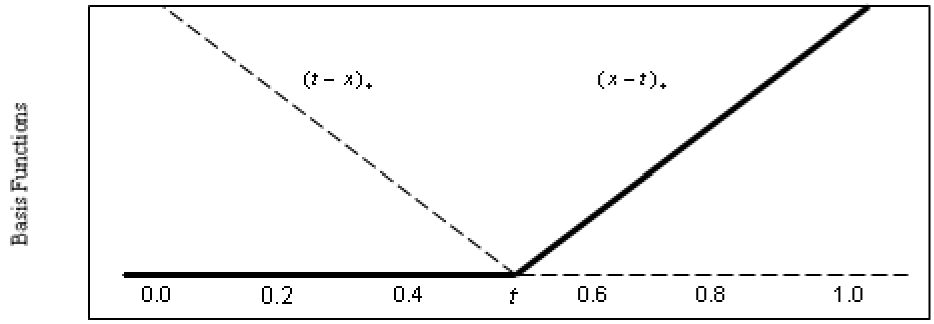
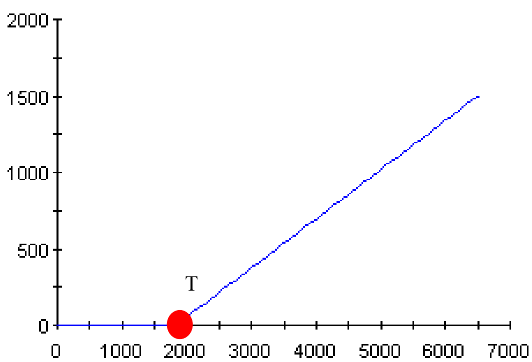
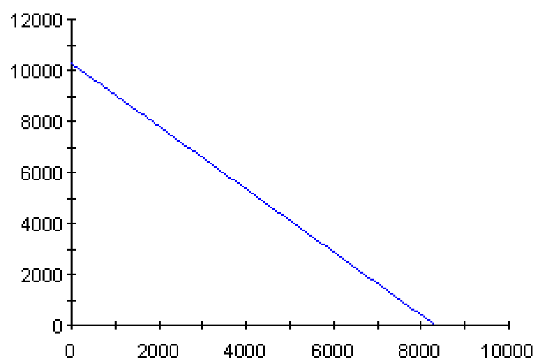


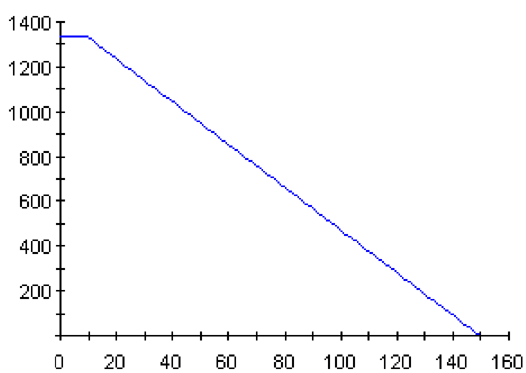
Figure 2: Illustration of Main Effect Shapes



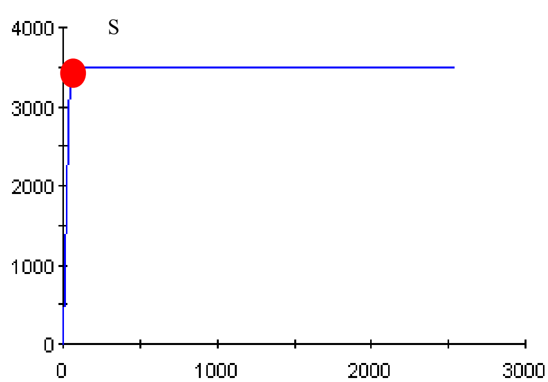
(a) Left Hockey-stick: Increasing after threshold. (Honda Civic Hybrid, Network TV)



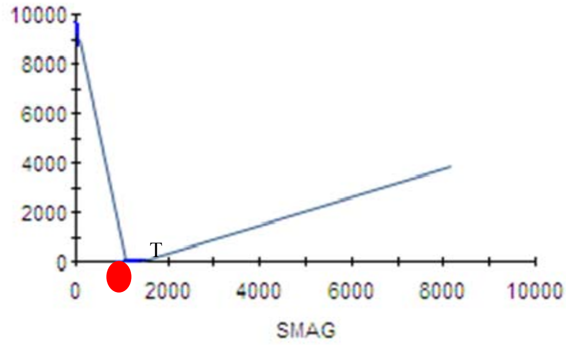
(b) Right Hockey-stick: Decreasing response- Inefficient allocation. (Jeep Grand Cherokee, Cable TV)



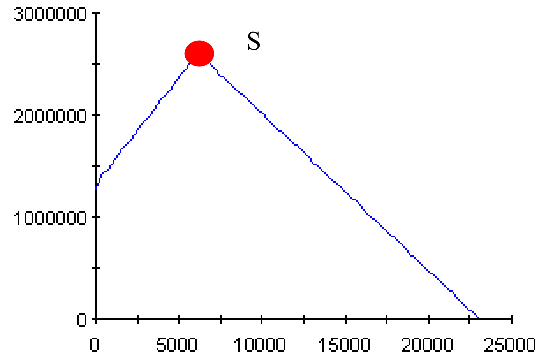
(c) Inverted Left Hockey-stick: Non-increasing with flat maximum response. Inefficient allocation. (Honda Civic Hybrid, Internet)



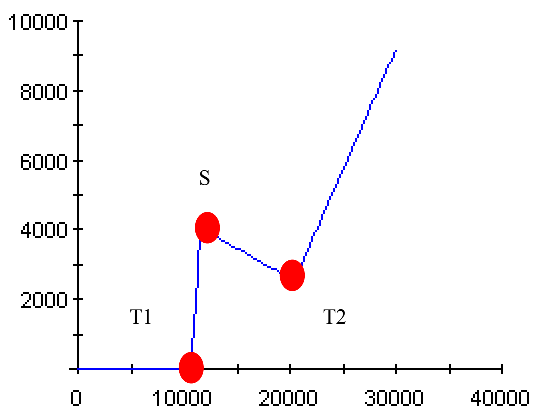
(d) Inverted Right Hockey-stick: Early saturation. (Ford Explorer, Newspapers)



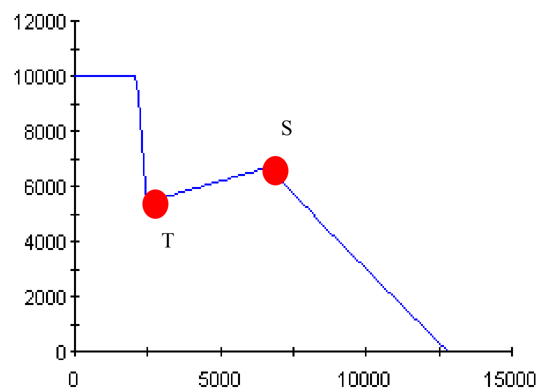
(e) V-shape: Nonmonotonic with threshold.
(Toyota Prius Hybrid, Magazines)



(f) Inverted-V: Nonmonotonic with super-saturation.
(Beer, Cable TV)

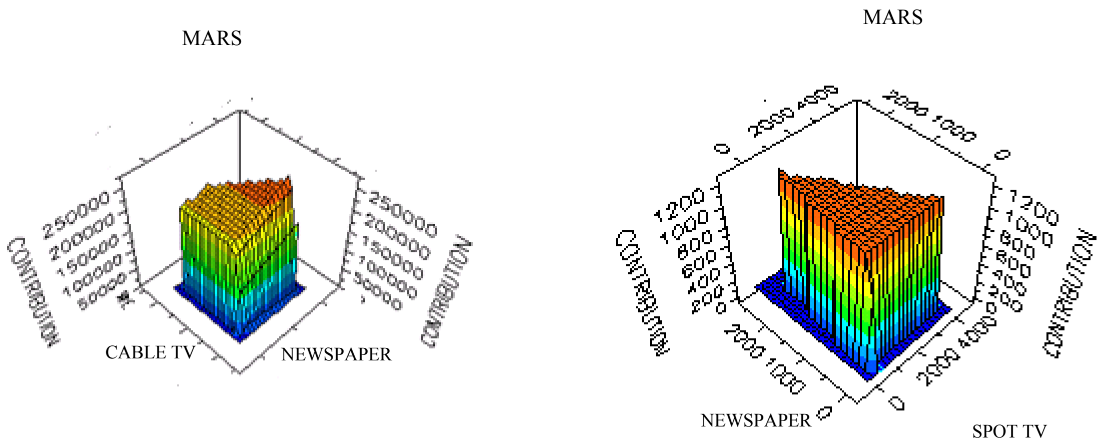


(g) Left Hockey-stick & V-shape combination: Nonmonotonic with double threshold and super-saturation.
(Jeep Grand Cherokee, Magazines)



(h) Inverted Left Hockey-stick & Inverted-V combination: Nonmonotonic with super-saturation and threshold.
(Ford Explorer, Network TV)

Figure 3: Nonmonotonic Media Interactions



(a) LNA **(b) Explorer**